

CRFE 1/21 03p

Access DB#

112340

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: David Fox Examiner #: 65461 Date: 1/15/04
 Art Unit: 1638 Phone Number: 202-0795 Serial Number: 09/936,869
 Mail Box and Bldg/Room Location: REM 2C18 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: 3/99

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

please do a SEQUENCE SEARCH

for : SEQ ID NO: 1 } promoter
 SEQ ID NO: 2 } variants of
 a single gene

1 NA 1320
 2 1401
 3 909
 5 1012

SEQ ID NO: 3 - protein encoded
 by gene associated
 with above
 promoters

and SEQ ID NO: 5 - coding sequence to
 be attached to
 above promoters

THANK YOU

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <u>4</u>	STN _____
Searcher Phone #: _____	AA Sequence (#) _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>1/21</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>1/22</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: <u>12</u>	Fulltext _____	Sequence Systems <u>af</u>
Clerical Prep Time: _____	Patent Family _____	WWW/Internet _____
Online Time: <u>12</u>	Other _____	Other (specify) _____

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 15:50:10 ; Search time 4813.93 Seconds
(without alignments)
11217.584 Million cell updates/sec

Title: US-09-936-869-1
Perfect score: 1320
Sequence: 1 aaactcgaacgaacaatatga.....tggtgctgttcgcccattg 1320

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
```

pred No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	275.8	20.9	1726	8	AF006210	Pinus res
2	201.8	15.3	1657	8	PM069482	U69482 Picea maria
3	170.6	12.9	1753	8	PADAL2	X79280 P.abies dal
4	146	11.1	1642	8	PMU64582	U64582 Picea maria
5	119.4	9.0	1555	8	PM069483	U69483 Picea maria
6	73.4	5.6	7218	6	I66494	I66494 Sequence 14
7	62.8	4.8	197271	2	AC121606	AC121606 Mus muscu
8	61.6	4.7	146895	2	AC119185	AC119185 Mus muscu
9	61	4.6	7218	6	I66494	I66494 Sequence 14
10	57.6	4.4	147708	2	AC129882	AC129882 Ornithorh
11	55.8	4.2	210677	2	AC074305	AC074305 Mus muscu
12	55	4.2	199033	2	AC134879	AC134879 Homo sapi
13	53.8	4.1	126611	10	AL928592	AL928592 Mouse DNA
14	53.2	4.0	181093	2	BX004853	BX004853 Dantio rer
15	53	4.0	210483	2	AC111508	AC111508 Rattus no
16	53	4.0	222431	2	AC106170	AC106170 Rattus no
17	53	4.0	245412	2	AC128428	AC128428 Rattus no
18	52	3.9	174929	2	AC134790	AC134790 Mus muscu
19	51.4	3.9	216905	2	AC091332	AC091332 Mus muscu
20	51.4	3.9	235902	2	AC031915	AC031915 Mus muscu
21	51.2	3.9	219342	10	AL732443	AL732443 Mouse DNA
22	51	3.9	208244	2	AC122545	AC122545 Mus muscu
23	51	3.9	231443	2	AC111050	AC111050 Mus muscu
24	50.8	3.8	213150	2	AC107642	AC107642 Mus muscu
25	50.8	3.8	223940	2	AC087567	AC087567 Mus muscu
26	50.6	3.8	246611	2	AC111404	AC111404 Rattus no
27	50.2	3.8	166648	2	AC116673	AC116673 Mus muscu
28	50.2	3.8	199996	2	AC102649	AC102649 Mus muscu
29	49.8	3.8	182649	2	AC023033	AC023033 Homo sapi
30	49.8	3.8	250029	3	AE014820	AE014820 Plasmodiu
31	49.8	3.8	252988	2	AC140457	AC140457 Mus muscu
32	49.8	3.8	260927	2	AC091771	AC091771 Mus muscu
33	49.6	3.8	29221	6	AX646443	AX646443 Sequence
34	49.6	3.8	29221	9	AB085699	AB085699 Homo sapi
35	49.6	3.8	29221	9	AC004976	AC004976 Homo sapi
36	49.6	3.8	127425	5	AL935063	AL935063 Zebrafish
37	49.6	3.8	141819	5	AL935063	AL935063 Mus muscu
38	49.6	3.8	176733	2	AC115006	AC115006 Mus muscu
39	49.6	3.8	219367	2	AC1137288	AC1137288 Mus muscu
40	49.6	3.8	285399	2	AC096348	AC096348 Rattus no
41	49.4	3.7	110000	2	BX119988	BX119988 Continuation (2 of
42	49.4	3.7	110000	2	BX119993	BX119993 Continuation (2 of
43	49.4	3.7	183853	9	AP003098	AP003098 Homo sapi
44	49.4	3.7	188498	2	AC009546	AC009546 Homo sapi
45	49.4	3.7	266530	2	AC127680	AC127680 Mus muscu
46	49.2	3.7	847	11	BV076803	BV076803 S212P602

ALIGNMENTS

RESULT 1	REFERENCE
AF006210	AUTHOR
LOCUS	TITLE
DEFINITION	JOURNAL
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

AF006210 1726 bp mRNA linear PLN 05-JAN-1999 complete cds.

Pinus resinosa
Pinus resinosus
 Eukaryota
 Eukaryotes
 Embryophyta; Embryophyta; Tracheophyta;
 Spermatophytes; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus:
 1 (base 1 to 1726)
 Liu, J.Y. and Podila, G.K.
 Direct Submission
 Submitted (01-JUN-1997) Department of Biological Sciences, Michigan

AUTHORS Rutledge, R.G., Kauffeldt, C., Fobert, P., Cote, C., Bosnich, W. and Stewart, D.

TITLE Direct Submission

JOURNAL Submitted (04-SEP-1996) Canadian Forest Service, Laurentian Forestry Centre, 1055 du P.E.P.S., Sainte-Foy, PQ G1V 4C7, Canada

ABSTRACTS Location/Qualifiers

SOURCE 1. 1657

```

/mol_type="mRNA"
/db_xref="taxon:3335"
/sex="female"
/tissue_type="cones"
/dev_stage="just prior to pollination"
457..1125

```

```

CDS
/dev_stages=just prior to polyomation
457..1125
/codon_start=1
/product="AGAMOUS-like MADS-box transcriptional factor
SAG1a"
/protein_id="AAC97157.1"
/db_xref="MCT:4033721"

```

```

SE COUNT      497 a 336 c 362 q 462 t
LQLG"
TCQENILASKTAEQCNQSHNTNMLSAFEYDALPFDSDNFEHLNLDAAHYAHQEQT
ILQVNSRGLRMDGDTALNILELQLEVRLEKGI GRVRSKKNNLLEEIDIMORRHEI
LIVANSHRGLFVGHANVKTIRYKXT CVDNHGGAI SESNSQYKQGLRQOIR
/translation="MGKGKIEIRIENTNTRQVTCRRNGLLKKAYELSVLDCDAEV
/GO_xref="GL:4033721"

```

SE COUNT	497 a	336 c	362 g	462 t
IGIN				

Query Match	15.3%;	Score 201.8;	DB 8;	Length 1657;
Best Local Similarity	88.7%;	Pred. No. 1.4e-38;		

Query Match 201.0; Score 201.0; Length 4057
 Best Local Similarity 88.7%; Pred. No. 1.4e-38;
 Matches 267; Conservative 0; Mismatches 22; Indels 12; Gaps 4

805 CGACCGACGACGCTCATCTTCTTCTCTGCGCAATTTTCTGCTGGATCATCATCAT 864

1 CGAACACCGACCGCTCATCTTCTCCTCTGCACA--TTTTCGTGCTGGATCATCATCAT 58

[illegible]

59 TATCATCATCGGCATCCCCACCACATCATCATGATAG-----TATCTCTCTTCGGCAG 112

925 TCGATTCTAGAGGAAAAGGAAGAGGAGAGGGGCAATATGATTTGTATCAACCTACCCGAAAAA 984

925 TCGATTGTAGAGGAACGGAAGAGGGGAGGGGCATATGTATTGATCAACTCCGAAAA 984
 113 TCAATCGTAGATGAAGCAAGAGGGGACGGGGCATA--TATTGATCAACTACCGAAAA 170

113 TCAATCGTAGTGAAGGAGGAGCGGGCATA--TATTGATCAACCTACCGAAAA 170
985 ACAATCTGATCAGCCCTGCTCAATCTTGGCTTATAATCTCTTATCCATGTTCATCAT 1048

985 ACAATCTGATCAGCCCTGCTCAATCTTGCTTATAAATCTCTTATCCACGTGTTCAATCAT 104
 |||||
 171 ACAACTGTATCAGCC--TGCTTAATCTGGCTTATAAAATCTCTTATCCACGTGTTCAATCAT 228
 |||||

171 ACAACCTGATCAGCC--TGCATATCTGGCTTATAAATCTTTATCCACCTGTTCAATCAAT 228
1045 CAGGTTTCTTCCCACTTTCAAGCAAAGGCCGCCGGATTTGGCCGTGTTCTTAGATTTTTCAG 1104

1045 CAGGTTTCTTCCACATTCAGCAAAAGCGCCGGATTGCCGTGTTCTTAGATTTCAG 110
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
229 CAGGTTTCTTCCCATTTGAGCAACAAGCGCCGGATTGCCGTGTTCTTAGATTTCAG 288
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

229 CAGGTTTCTTCCCACTTTGAGCAACGGCGCCGGATTGGCCGTGTTCTTAGATTTTAG 288

1105 G 1105

1105 G 1105
289 G 289

289 G 289

SULT 3
DAL2
CIV

DAL2	CUS	PADAL2	1753 bp	mRNA	linear	PLN 28-FEB-1995
		P.abies dal2 mRNA.				
		VFNTION				
		VF0280				
		VF0280				

FINITION P.abies dal2 mRNA.
CESSION X79280
X79280.1 GI:695687
DAL2 gene
X79280

X79280.1 GI:695687
 DAL2 gene.
 Picea abies (Norway spruce)
 URCE
 ORGANISM Picea abies

ORGANISM

Picea abies (Norway spruce)

Picea abies

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.

REFERENCE
AUTHORS
Tandre K., Albert V.A., Sundas, A. and Engstrom, P.
1
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta

1
REFERENCE
AUTHORS
Tandre, K., Albert, V.A., Sundas, A. and Engstrom, P.
TITLE
Conifer homologues to genes that control floral development in
angiosperms

angiosperms
Plant Mol. Biol. 27 (1), 69-78 (1995)
MEDLINE 9517009

WEDLINE 95170009

Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big_Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 197248 bases at least Q40
Consensus quality: 197263 bases at least Q30
Consensus quality: 197264 bases at least Q20
Insert size: 212000; agarose-fp
Insert size: 19983; sum-of-contigs
Quality coverage: 14.60 in Q20 bases; agarose-fp
Quality coverage: 12.64 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

REFERENCE
AUTHORS

2 (bases 1 to 146695)

Birren, B., Linton, L., Nusbbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boughalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B., Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArallano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lechoczky, J., Levine, R., Lindblad-Rohr, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihovata, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Ratta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, J., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggilo, J., Vassiliou, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A., and Zody, M.

TITLE
JOURNAL

COMMENT

Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 147708)
 Green, E.D.
 Direct Submission
 Submitted (06-AUG-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 147708)
 Green, E.D.
 Direct Submission
 Submitted (23-AUG-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Aug 23, 2002 this sequence version replaced gi:22122885.

 Genomic Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.zooenhgri.nih.gov

 Project Information
 Center project name: dwc
 Center clone name: 265D21

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 146677 bases at least Q40
 Consensus quality: 147063 bases at least Q30
 Consensus quality: 147215 bases at least Q20
 Insert size: 148000; agarose-fp
 Insert size: 147408; sum-of-contigs
 Quality coverage: 9.31x in Q20 bases; agarose-fp
 Quality coverage: 9.35x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 87687: contig of 87687 bp in length
 87688 87687: gap of unknown length
 87788 90934: contig of 3147 bp in length
 90935 91034: gap of unknown length
 91035 124474: contig of 33440 bp in length
 124475 124574: gap of unknown length
 124575 147708: contig of 23134 bp in length.

FEATURES

1..147708
 /organism="Ornithorhynchus anatinus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9258"
 /clone="CLM1-265D21"
 /clone_lib="CLM1"

1..87687
 /note="assembly_fragment
 clone_end:17"

misc_feature
 1..34612
 /note="clone overlaps with GenBank Accession Number
 AC130185 clone CLM1-379P1 (center project name dwd)"
 74481..147708
 /note="clone overlaps with GenBank Accession Number
 AC127463 clone CLM1-27L6 (center project name dwd)"
 misc_feature
 87788..90934
 /note="assembly_fragment"
 misc_feature
 91035..124474
 /note="assembly_fragment"
 misc_feature
 124575..147708
 /note="assembly_fragment
 clone_end:SP6
 vector_side:right"
 vector_side:left"

BASE COUNT 40144 a 33940 c 33579 g 39745 t 300 others
 ORIGIN

Query Match 4.4%; Score 57.6; DB 2; Length 147708;
 Best Local Similarity 50.0%; Pred. No. 0.0027;
 Matches 144; Conservative 0; Mismatches 144; Indels 0; Gaps 0;
 QY 226 TGTCTGGGGCTTGGATGGCTTGGCTTGGTGGAGAGCCCAACAAATTTATAGAAATATATA 285
 Db 88401 TGTCTGGGGCTTGGGGTAAATGGAATGTCTGAGAGAGAGAGAGAGAGAGAGAGAG 88342
 QY 286 AATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 345
 Db 88341 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 88282
 QY 346 AGAAGAGTACTTTAACAATGCGCAACCAAGGAGAGATTCACAACTTCTGACCTCG 405
 Db 88281 AG 88222
 QY 406 AATACAGATATATGGTGGTAAGAAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 465
 Db 88221 AGAAG 88162
 QY 466 AGAGCCACAAATGGAG 513
 Db 88161 AGAAG 88114

RESULT 11

AC074305

LOCUS

DEFINITION

AC074305

AC074305

AC074305.1

GI:9454521

HTG; HTGS_PHASE1; HTGS_DRAFT.

Mus musculus (house mouse)

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 210677)

Sequencing of Mouse

Unpublished

2 (bases 1 to 210677)

DOE Joint Genome Institute.

Direct Submission

Submitted (26-JUL-2000) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: http://www.jgi.doe.gov

Project Information

Center Project Name: 0

Center clone name: RPC1-23_1012

AUTHORS
 TITLE
 JOURNAL
 Waterston, R.H.
 Direct Submission
 Submitted (01-OCT-2002) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 3 (bases 1 to 199033)
 Waterston, R.H.
 Direct Submission
 Submitted (23-FEB-2003) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On Feb 23, 2003 this sequence version replaced gi:23396428.

COMMENT

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 Contact: submissions@watson.wustl.edu
 ----- Project Information -----
 Center project name: H.NH0295P22

----- Summary Statistics -----
 Sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: Dye-terminator; 0% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 193854 bases at least Q40
 Consensus quality: 195124 bases at least Q30
 Consensus quality: 195874 bases at least Q20

* NOTE: This is a "working draft" sequence. It currently
 * consists of 13 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 2110: contig of 2110 bp in length.
 * 2111 2210: gap of unknown length
 * 2211 4639: contig of 2429 bp in length
 * 4640 4739: gap of unknown length
 * 4740 10008: contig of 5269 bp in length
 * 10009 10108: gap of unknown length
 * 10109 13836: contig of 3728 bp in length
 * 13837 13936: gap of unknown length
 * 13937 17705: contig of 3769 bp in length
 * 17706 17805: gap of unknown length
 * 17806 24927: contig of 7122 bp in length
 * 24928 25027: gap of unknown length
 * 25028 31944: contig of 6917 bp in length
 * 31945 32044: gap of unknown length
 * 32045 50505: contig of 18461 bp in length
 * 50506 50606: gap of unknown length
 * 50606 63617: contig of 13012 bp in length
 * 63617 82874: contig of 19157 bp in length
 * 82875 82975: gap of unknown length
 * 82975 106052: contig of 23078 bp in length
 * 106053 131178: gap of unknown length
 * 106153 131178: contig of 25026 bp in length
 * 131179 131279: gap of unknown length
 * 131279 199033: contig of 67755 bp in length.

FEATURES
 source

1. 199033
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="UNK"
 /clone="RP11-295P22"
 1. 2110
 /note="assembly_name:Contig17"

misc_feature

misc_feature 2211..4639
 /note="assembly_name:Contig18"
 misc_feature 4740..10008
 /note="assembly_name:Contig19"
 misc_feature 10109..13836
 /note="assembly_name:Contig20"
 misc_feature 13937..17705
 /note="assembly_name:Contig21"
 misc_feature 17806..24927
 /note="assembly_name:Contig22"
 misc_feature 25028..31944
 /note="assembly_name:Contig23"
 misc_feature 32045..50505
 /note="assembly_name:Contig24"
 misc_feature 50606..63617
 /note="assembly_name:Contig25"
 misc_feature 63718..82874
 /note="assembly_name:Contig26"
 misc_feature 82975..106052
 /note="assembly_name:Contig27"
 misc_feature 106153..131178
 /note="assembly_name:Contig28"
 misc_feature 131279..199033
 /note="assembly_name:Contig29"
 BASE COUNT 57536 a 42075 c 35735 g 62476 t 1211 others
 ORIGIN
 Query Match 4.2%; Score 55; DB 2; Length 199033;
 Best Local Similarity 58.1%; Pred. No. 0.012;
 Matches 97; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
 QY 178 AAGAAGTAAAGGATTGACTCTAGTCAAGTCAAGTGAAGGATTGGAAGGACTTG 237
 Db 154397 AATGGAATCAATGATGGAATCAAGTGAAGGATTGGAAGGACTTG 154338
 QY 238 GATGGCTGGGTTGCTGTGAGAAGCCACAAATTTATAGAAATATATAAAAAATA 297
 Db 154337 AATGATTTGGGTAGAGAAGAAATGGAATCTATCGGTGGAATAGAAATAGAAAGAA 154278
 QY 298 AAAAAATTAAGTGTGGAAAGTGAAGGAAACCGTGGGCGAAGAAATATACA 344
 Db 154277 AAAGAATTTAATCGAGTGTGATGGAACATATGGAATGGAATTGAAA 154231
 RESULT 13
 AL928592
 LOCUS
 DEFINITION Mouse DNA sequence from clone RP24-168J7 on chromosome 4, complete
 sequence.
 ACCESSION AL928592
 VERSION AL928592.13 GI:31071444
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 126611)
 Giselle H.
 Direct Submission
 Submitted (22-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On May 23, 2003 this sequence version replaced gi:29125348.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to confirm this sequence. Sequence data
 from the whole genome shotgun alone has only been used where it has
 a phred quality of at least 30.
 ----- Genome Center -----
 Center: UK Medical Research Council
 Center code: UK-MRC
 Web site: <http://mrcseq.har.mrc.ac.uk>
 Contact: mousedq@har.mrc.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emu, EMBL; Swi, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP24-168J7 is from a Male (C57BL/6J) mouse BAC Library VECTOR: pTARBAC1.

FEATURES
Source
1..126611
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP24-168J7"
/clone_lib="RPCI-24"
BASE COUNT 40625 a 24197 c 24348 g 37441 t
ORIGIN

Query Match 4.1%; Score 53.8; DB 10; Length 126611;
Best Local Similarity 49.8%; Pred. No. 0.023;
Matches 136; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 230 GGGCGCTGGATGCTTGGCTTGGTGTGAGAACCAACATTTATAGAAATATATATAAT 289
Db 31351 GTGGTGGGACGAGGAGGAGAGTATAGAGGTTTATGGATAGCATTTGAAT 31410

QY 290 AAAAAATAAAAATTTAGTGTGCACTGAACGCTGGGGGAGAAATATACACGAA 349
Db 31411 GTAAATGAAAAATATTTTAAATAAATTTTCAAAAGAAAGAAAGAAAGAA 31470

QY 350 GAGTACTTTTAAACATGCGCAACCAAGGACAGATTCACACTTGTGACCTCGAATA 409
Db 31471 AGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 31530

QY 410 CGAGATATGCTGTGTAAGAAATTAAGAGAGAGTGGAGTGCATTTGAAATGAGTGGAG 469
Db 31531 AGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 31590

QY 470 CGCACAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 502
Db 31591 AGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 31623

RESULT 14
BX004853
LOCUS
DEFINITION
Dario rerio clone DKEY-13011, *** SEQUENCING IN PROGRESS ***, 2
unordered pieces.
ACCESSION
BX004853
VERSION
BX004853.5
KEYWORDS
HTG; HTGS PHASE1; HTGS ACTIVEFIN; HTGS DRAFT; HTGS FULLTOP.
SOURCE
Dario rerio
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 181093)
REFERENCE
Harrison, E.
Direct Submission
JOURNAL
Submitted (02-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,

COMMENT

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-hel@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 3, 2003 this sequence version replaced gi:30024305.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfsh-hel@sanger.ac.uk
----- Project Information
Center project name: zk13011
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 180523 bases at least Q40
Consensus quality: 180751 bases at least Q30
Consensus quality: 180905 bases at least Q20
Insert size: 180993; sum-of-contigs
Insert size: 172880; 4.9% error; agarose-fp
Quality coverage: 6.77x in Q20 bases; sum-of-contigs Quality
coverage: 7.52x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 95025: contig of 95025 bp in length
* 95026 95125: gap of 100 bp
* 95126 181093: contig of 85968 bp in length.

FEATURES
Source

1..181093
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-13011"
/clone_lib="DanioKey"

misc_feature
1..95025
/note="assembly_fragment:00925"

misc_feature
95126..181093
/note="assembly_fragment:02465.0"

BASE COUNT 56902 a 34623 c 33889 g 55579 t 100 others
ORIGIN

Query Match 4.0%; Score 53.2; DB 2; Length 181093;
Best Local Similarity 51.2%; Pred. No. 0.032;
Matches 124; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 257 AGAAGCCAACTTTTATAGAAATATATATAATAAAAAATTTAAGTTTGA 316
Db 81583 AAAAGCAATCTAAGGAAAAATCAATGGAACAAAGAAATATAAATTAACGTAA 81642

QY 317 AGTGAACCGTGGGCGAGAAATATACACAGACAGAGTACTTTTAACTGCA 376
Db 81643 AATGAAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 81702

QY 377 CAGATTCAACTTGAATTTCTGGACCTCGAATACAGATATATGTTGTAAGAAATAAGG 436
Db 81703 CAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 81762

QY 437 AAGCTGGAGTGCATTTGAAATGATGAGGCGCAGAAATGAGACGAAATAATGA 496
Db 81763 AAGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 81822

QY 497 AA 498
Db 81823 AA 81824

RESULT 15
AC111508/c

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 15:50:10 ; Search time 5109.33 Seconds
(without alignments)
11217.584 Million cell updates/sec

Title: US-09-936-869-2
Perfect score: 1401
Sequence: 1 aaactgcacagcaaatatga.....ggaaattttgtacaaatc 1401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.on.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sv.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.on.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.un.*
- 28: em.vi.*
- 29: em.htg.hum.*
- 30: em.htg.inv.*
- 31: em.htg.other.*
- 32: em.htg.mus.*
- 33: em.htg.pln.*
- 34: em.htg.rod.*
- 35: em.htg.mam.*
- 36: em.htg.vrt.*
- 37: em.sv.*
- 38: em.htgo.hum.*
- 39: em.htgo.mus.*
- 40: em.htgo.pln.*
- 41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	280	20.0	1726	8	AF006210	Pinus res
2	201.8	14.4	1657	8	PMU69482	U69482 Picea maria
3	170.6	12.2	1753	8	PADAL2	X79280 P-abies dal
4	146	10.4	1642	8	PMU46582	U4582 Picea maria
5	119.4	8.5	1555	8	PMU69483	U69483 Picea maria
6	73.4	5.2	7218	6	I66494	I66494 Sequence 14
7	62.8	4.5	197271	2	AC121606	AC121606 Mus muscu
8	61.6	4.4	146695	2	AC119185	AC119185 Mus muscu
9	61	4.4	7218	6	I66494	I66494 Sequence 14
10	57.6	4.1	147708	2	AC129882	AC129882 Ornithoch
11	55.8	4.0	210677	2	AC074305	AC074305 Mus muscu
12	55	3.9	199033	2	AC134879	AC134879 Homo sapi
13	53.8	3.8	126611	10	AL928592	AL928592 Mouse DNA
14	53.2	3.8	181093	2	BX004853	BX004853 Danio rer
15	53	3.8	210483	2	AC111508	AC111508 Rattus no
16	53	3.8	222431	2	AC106170	AC106170 Rattus no
17	53	3.8	245412	2	AC128428	AC128428 Rattus no
18	52	3.7	174929	2	AC134790	AC134790 Mus muscu
19	51.4	3.7	216905	2	AC091332	AC091332 Mus muscu
20	51.4	3.7	235902	2	AC139195	AC139195 Mus muscu
21	51.2	3.7	219342	10	AL732443	AL732443 Mouse DNA
22	51	3.6	208244	2	AC122545	AC122545 Mus muscu
23	51	3.6	231443	2	AC111050	AC111050 Mus muscu
24	50.8	3.6	213150	2	AC107642	AC107642 Mus muscu
25	50.8	3.6	223940	2	AC087567	AC087567 Mus muscu
26	50.6	3.6	246611	2	AC111404	AC111404 Rattus no
27	50.2	3.6	166489	2	AC116673	AC116673 Mus muscu
28	50.2	3.6	199996	2	AC102649	AC102649 Mus muscu
29	49.8	3.6	182649	2	AC023033	AC023033 Homo sapi
30	49.8	3.6	250029	3	AE014820	AE014820 Plasmodiu
31	49.8	3.6	252888	2	AC140457	AC140457 Mus muscu
32	49.8	3.6	260967	2	AC091771	AC091771 Mus muscu
33	49.6	3.5	29221	9	AX646443	AX646443 Sequence
34	49.6	3.5	29221	9	AB065699	AB065699 Homo sapi
35	49.6	3.5	127425	9	AC004976	AC004976 Homo sapi
36	49.6	3.5	141819	5	AL935063	AL935063 Zebraphish
37	49.6	3.5	176733	2	AC115006	AC115006 Mus muscu
38	49.6	3.5	219367	2	AC127288	AC127288 Mus muscu
39	49.6	3.5	285389	2	AC096348	AC096348 Rattus no
40	49.4	3.5	110000	2	EX119988_1	Continuation (2 of
41	49.4	3.5	110000	2	EX119993_1	Continuation (2 of
42	49.4	3.5	183853	9	AP003098	AP003098 Homo sapi
43	49.4	3.5	188498	2	AC009546	AC009546 Homo sapi
44	49.4	3.5	266530	2	AC127680	AC127680 Mus muscu
45	49.2	3.5	847	11	BV076803	BV076803 S212P6027

ALIGNMENTS

RESULT 1

AF006210 AF006210 1726 bp mRNA linear PLN 05-JAN-1999
LOCUS Pinus resinosa MADS box transcription factor mRNA, complete cds.
DEFINITION Pinus resinosa MADS box transcription factor mRNA, complete cds.
ACCESSION AF006210
VERSION AF006210.1 GI:4101709

KEYWORDS

SOURCE

ORGANISM

Pinus resinosa

Pinus resinosa

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

1 (bases 1 to 1726)

Liou, J.-J. and Podila, G.K.

Direct Submission

TITLE

JOURNAL

Pred. No. is the number of results predicted by chance to have a

Db 289 G 289

RESULT 3
PADAL2
LOCUS
DEFINITION
P.abies dal2 mRNA.
X79280
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

PADAL2
P.abies dal2 mRNA.
X79280
GI:695687
DAL2 gene.
Picea abies (Norway spruce)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Pinaceae; Picea.

Tandre,K., Albert,V.A., Sundas,A. and Engstrom,P.
Conifer homologues to genes that control floral development in
angiosperms
Plant Mol. Biol. 27 (1), 59-78 (1995)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
DIRECT Submission
Submitted (16-MAY-1994) K. Tandre, Uppsala University, Dept of
Physiological Botany, Villavagen 5, 752 36 Uppsala, SWEDEN
Location/Qualifiers
1..1753
/organism="Picea abies"
/mol_type="mRNA"
/db_xref="taxon:3329"
/clone="3ae"
/tissue_type="female cone"
/clone_lib="cDNA library spruce female; SF"
/dev_stage="before formation of ovule"
1..1753
/gene="dal2"
424..1092
/gene="dal2"
/product="DAL2 protein"
protein_id="CAA5867.1"
/db_xref="GI:695688"
/db_xref="SPTREMBL:Q40766"
translation="MGRGKIEIKRIENTNRVTFCKRRNGLLKAYELSVLCDAVA
LIVFSRRLVEFANHSVKRTIERYKTCDVNHHGGVISNSQYWOAEAGKLQQDIL
IQENLRSLAKIAEQCNHNTWLSAPDYALPAFDSRNFLHANLIIDAAHYHQEQTT
LQLG"

misc_feature
427..594
/gene="dal2"
note="MADS box"

misc_feature
597..891
/gene="dal2"
note="K box"

BASE COUNT 529 a 340 c 388 g 496 t
ORIGIN

Query Match 12.2%; Score 170.6; DB 8; Length 1753;
Best Local Similarity 88.6%; Pred.No. 4;Se=31;
Matches 233; Conservative 0; Mismatches 15; Indels 11; Gaps 4;

QY 843 TTTCCTGTGGATCATCATATTACCATCCGCCATCCCCCACCATCATCATCATGATGG 902
DB 5 TTTCCTGTGGATCATCATATTACCATCCGCCATCCCC-CCATCATCATCATGATG 63

QY 903 TATCTTATCTCTCCCTGCATCGATTGTAGAGAGAAGGAAGGAAGGCGGATATGT 962
DB 64 -----TATCTCTCTTGCGAGTCAATCGTAGAAGAAGGAAGAACGCGGGATA--T 115

QY 963 ATTGATCAACTACCGAAAAACAATCTGATCAGCCCTGCTCAATCTTGTTTAATC 1022
116 ATTGATCAACTACCGAAAAACAATCTGATCAGCC---TGCTATCTGGTTATAATC 173

QY 1023 TCCTATCCACTGTTCATCATCTTCCACATTTTCCACATTTCAAGCAAGACGCGCCGATT 108
DB 174 TCCTATCCACTGTTCATCATCTTCCAGTTTTCTTCCACTTTTGAACGACGCGCCGATT 233

QY 1083 GCCTGTGTCTTAGATTTTCAGG 1105
DB 234 GCCTGTGTCTTAGATTTTAGG 256

PMU46582 1642 bp mRNA linear PLN 18-DEC-1998
Picea mariana AGAMOUS-like MADS-box transcription factor SNADS42B
mRNA, complete cds.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

PMU46582
U46582
U46582.1 GI:4033709
Picea mariana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Pinaceae; Picea.
1 (bases 1 to 1642)
Rutledge,R., Regan,S., Nicolas,C., Fobert,P., Cote,C., Bosnich,W.,
Kauffeldt,C., Sunohara,G., Seguin,A. and Stewart,D.
Characterization of an AGAMOUS homologue from the conifer black
spruce (*Picea mariana*) that produces floral homeotic conversions
when expressed in Arabidopsis
Plant J. 15 (5), 625-634 (1998)

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
DIRECT Submission
Submitted (17-JAN-1996) Robert G. Rutledge, Molecular Genetics and
Tissue Culture, Petawawa National Forestry Institute, Chalk River,
ON K0J 1J0, Canada
Location/Qualifiers
1..1642
/organism="Picea mariana"
/mol_type="mRNA"
/db_xref="taxon:3335"
/sex="female"
/tissue_type="cones"
/dev_stage="just prior to pollination"
390..1058
/codon_start=1
/product="AGAMOUS-like MADS-box transcription factor
SNADS42B"
protein_id="AAC97146.1"
/db_xref="GI:4033710"

CDS
base count start=1
SMADS42B

BASE COUNT 485 a 321 c 366 g 470 t
ORIGIN

Query Match 10.4%; Score 146; DB 8; Length 1642;
Best Local Similarity 86.4%; Pred.No. 5e+25;
Matches 185; Conservative 0; Mismatches 25; Indels 4; Gaps 2;

QY 892 CATCATGATGGTATCTTCTTCTCCCTGGCAATCGATTGTAGAGAGAAGGAAGGAAA 951
DB 13 CATCATCATCATGATGATGATCTTCTTCCGAGTCATCGTAGAGAAAGGAAGAGAC 72

QY 952 GGCGATATGATTTCATCACTACCTACCGAAAAACAATCTGATCAGCCCTGCTCAATCTT 1011
DB 73 GGCGATA--TATTCATCACTACCGAAAAACAATCTGATCAGCC--TGCTAATCTG 128

QY 1012 GTTTAATCTTCTTATTCATCTTCAATCATCTAGTTCTTCTTCCCATCTTTCAGCAAG 1071

RESULT 6	7
LOCUS	LOCUS
166494/c	166494/c
DEFINITION	DEFINITION
Sequence 14 from patent US 5670367.	Sequence 14 from patent US 5670367.
ACCESSION	ACCESSION
166494	166494
VERSION	VERSION
166494.1	166494.1
KEYWORDS	KEYWORDS
Unknown.	Unknown.
SOURCE	SOURCE
Unknown.	Unknown.
ORGANISM	ORGANISM
Unclassified.	Unclassified.
REFERENCE	REFERENCE
1 (bases 1 to 7218)	1 (bases 1 to 7218)
AUTHORS	AUTHORS
Dorner, F., Scheiflinger, F. and Falkner, F. Gunter.	Dorner, F., Scheiflinger, F. and Falkner, F. Gunter.
TITLE	TITLE
Recombinant fowlpox virus	Recombinant fowlpox virus
JOURNAL	JOURNAL
Patent: US 5670367-A 14 23-SEP-1997;	Patent: US 5670367-A 14 23-SEP-1997;
FEATURES	FEATURES
Location/Qualifiers	Location/Qualifiers
1..7218	1..7218
source	source
/organism="unknown"	/organism="unknown"
BASE COUNT	BASE COUNT
1944 a 1491 c 1486 g	1929 t 368 others
ORIGIN	ORIGIN
Query Match	Query Match
Best Local Similarity	Best Local Similarity
6.2%; Score 73.4; DB 6; Length 7218;	6.2%; Score 73.4; DB 6; Length 7218;
Matches 26; Conservative 236; Mismatches 157; Indels 0; Gaps 0;	Matches 26; Conservative 236; Mismatches 157; Indels 0; Gaps 0;
QY	QY
93 CACTGTACTTTGAGAAAGGTAGTGGACTCTGCTCAGGTACATTAGTTTGGTAAGTGGC 152	93 CACTGTACTTTGAGAAAGGTAGTGGACTCTGCTCAGGTACATTAGTTTGGTAAGTGGC 152
DB	DB
1487 CACTGTAAATTCCTATCTATGCAAGTAGTTAAAGAGATAGAAGAAATTTGGTACR 1428	1487 CACTGTAAATTCCTATCTATGCAAGTAGTTAAAGAGATAGAAGAAATTTGGTACR 1428
QY	QY
153 TTGCGTCTCGGTAAATATGAGAAGTAAGAAGTAAAGATTTGACTCTAGTCAAGTAC 212	153 TTGCGTCTCGGTAAATATGAGAAGTAAGAAGTAAAGATTTGACTCTAGTCAAGTAC 212
DB	DB
1427 RRR 1368	1427 RRR 1368
QY	QY
213 ATTGGATTGCTTTGTGGGGCTTGATGGCTTGGGTTGGTGTGAGAACCAACAATT 272	213 ATTGGATTGCTTTGTGGGGCTTGATGGCTTGGGTTGGTGTGAGAACCAACAATT 272
DB	DB
1367 RRR 1308	1367 RRR 1308
QY	QY
273 TAAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 332	273 TAAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 332
DB	DB
1307 RRR 1248	1307 RRR 1248
QY	QY
333 CAGAAATATACAGAGAAGTACTTTAAACAATGCGCAACCAAGCAGATTCACAATTGA 392	333 CAGAAATATACAGAGAAGTACTTTAAACAATGCGCAACCAAGCAGATTCACAATTGA 392
DB	DB
1247 RRR 1198	1247 RRR 1198
QY	QY
393 TTTTGGACCTCGAATACAGAGATAATGCTGCTGAAGAAATAAGAGAGAGTGGAGTGC 452	393 TTTTGGACCTCGAATACAGAGATAATGCTGCTGAAGAAATAAGAGAGAGTGGAGTGC 452
DB	DB
1187 RRR 1128	1187 RRR 1128
QY	QY
453 TGAAATGAATGGAGACGCACAAAATGGAGGACGAATAAATGAATAATATGCAAGAG 511	453 TGAAATGAATGGAGACGCACAAAATGGAGGACGAATAAATGAATAATATGCAAGAG 511
DB	DB
1127 RRR 1069	1127 RRR 1069
RESULT 7	RESULT 7
AC121606/c	AC121606/c
LOCUS	LOCUS
197271 bp	197271 bp
DEFINITION	DEFINITION
Mus musculus chromosome UNK clone RP23-314HS, WORKING DRAFT	Mus musculus chromosome UNK clone RP23-314HS, WORKING DRAFT
SEQUENCE.	SEQUENCE.
AC121606	AC121606
AC121606.2	AC121606.2
GI:22475583	GI:22475583
VERSION	VERSION
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ACTIVEFIN.	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ACTIVEFIN.
KEYWORDS	KEYWORDS
Mus musculus (house mouse)	Mus musculus (house mouse)
SOURCE	SOURCE
Mus musculus	Mus musculus
ORGANISM	ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	REFERENCE
1 (bases 1 to 197271)	1 (bases 1 to 197271)
AUTHORS	AUTHORS
McPherson, J.D. and Waterston, R.H.	McPherson, J.D. and Waterston, R.H.
TITLE	TITLE
The sequence of Mus musculus clone	The sequence of Mus musculus clone
JOURNAL	JOURNAL
Unpublished	Unpublished
REFERENCE	REFERENCE
2 (bases 1 to 197271)	2 (bases 1 to 197271)
AUTHORS	AUTHORS
McPherson, J.D. and Waterston, R.H.	McPherson, J.D. and Waterston, R.H.
TITLE	TITLE
Direct Submission	Direct Submission

```

JOURNAL Submitted (20-MAY-2002) Genome Sequencing Center, 4444 Forest Park
REFERENCE Parkway, St. Louis, MO 63108, USA
AUTHORS 3 (bases 1 to 197271)
McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Aug 25, 2002 this sequence version replaced gi:20986693.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_EA0314H05
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-terminator Big Dye; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 197248 bases at least Q40
Consensus quality: 197263 bases at least Q30
Consensus quality: 197264 bases at least Q20
Insert size: 212000; agarose-fp
Insert size: 199833; sum-of-contigs
Quality coverage: 14.60 in Q20 bases; agarose-fp
Quality coverage: 12.64 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 197271: contig of 197271 bp in length.
-----
FEATURES
Location/Qualifiers
1..197271
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-314H5"
misc_feature 1..197271
/note="assembly name:Contig12" 4 others
BASE COUNT 56926 a 43271 c 42141 g 54929 t
ORIGIN
Query Match 4.5%; Score 62.8; DB 2; Length 197271;
Best Local Similarity 52.7%; Pred. No. 0.00014;
Matches 136; Conservative 0; Mismatches 122; Indels 0; Gaps 0;
QY 245 TGGGTTTCGTGAGAACCAATTTATAGAATATATAATAATAATAATAATAATAATA 304
Db 152238 TGGCTTTTGTGTAATAAGGAAGACATTTGTACTTTAAANAAAAAAGAAAAA 152179
QY 305 TTAAGTTTGGAAAGTGAACCGTGGGCGCAAAATATACACAGAAGACTTTTACAT 364
Db 152178 TCTTGTCTCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA 152119
QY 365 GCGCAACCAAGCGCAATTCAACCTTGATTTCTCGACCTCGAATACGAGATATCGTGT 424
Db 152118 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA 152059
QY 425 AAGCAATTAAGCAAGAGTGGAGTGTCATTTGAAAAATGAATGAGAGCGCAAAATGAGG 484
Db 152058 AAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGGGAAGAAGAAGAGAGA 151999
QY 485 ACGAATAAATGAAATATA 502

```

Db 151998 AGRAGAGAGGAGAGAA 151981

RESULT 8
AC119185
LOCUS
DEFINITION

Mus musculus clone RP24-174M19, WORKING DRAFT SEQUENCE, 7 unordered pieces.
AC119185 146695 bp DNA linear HTG 07-MAR-2003
HTG; HTGS_PHASE1; HTGS_DRAFT.
Mus musculus (house mouse)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
1 (bases 1 to 146695)
Birren,B., Nusbaum,C. and Lander,E.
Mus musculus, clone RP24-174M19
Unpublished
2 (bases 1 to 146695)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,N., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Lamazares,R.,
Kamat,A., Karatas,A., Kells,C., Laocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.N.,
Seaman,S., Severy,P., Spencer,B., Stange-Thmann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 146695)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Bouckgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hakez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Landblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,J., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Spencer,B., Stange-Thmann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (07-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 7, 2003 this sequence version replaced gi:20303909.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L25189
 Center clone name: 174_M19

----- Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 144877 bases at least Q40
 Consensus quality: 145590 bases at least Q30
 Consensus quality: 145872 bases at least Q20
 Insert size: 145000; agarose-fp
 Insert size: 146095; sum-of-contigs
 Quality coverage: 8.6 in Q20 bases; agarose-fp
 Quality coverage: 8.5 in Q20 bases; sum-of-contigs

***** NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 5643: contig of 5643 bp in length
* 5644 5743: gap of 100 bp
* 5744 75707: contig of 69964 bp in length
* 75708 75807: gap of 100 bp
* 75808 82448: contig of 6841 bp in length
* 82449 82549: gap of 100 bp
* 82549 97703: contig of 15155 bp in length
* 97704 97804: gap of 100 bp
* 97804 115886: contig of 18083 bp in length
* 115887 115987: gap of 100 bp
* 115987 144141: contig of 28155 bp in length
* 144142 144242: gap of 100 bp
* 144242 146695: contig of 2454 bp in length.

```

FEATURES

source

```

1..146695
  /organism="Mus musculus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10090"
  /clone="RP24-174M19"
  /clone_lib="RPCL-24 Male Mouse BAC"

```

misc_feature

```

1..5643
  /note="assembly_fragment"
  clone_end:896
  vector_side:left

```

misc_feature

```

5744..75707
  /note="assembly_fragment"

```

misc_feature

```

75808..82448
  /note="assembly_fragment"

```

misc_feature

```

82549..97703
  /note="assembly_fragment"

```

misc_feature

```

97804..115886
  /note="assembly_fragment"

```

misc_feature

```

115987..144141
  /note="assembly_fragment"

```

misc_feature

```

144242..146695
  /note="assembly_fragment"
  clone_end:7

```

vector_side:right

BASE COUNT 42577 a 32450 c 32243 g 38825 t 600 others

ORIGIN

Query Match 4.4%; Score 61.6; DB 2; Length 146695;
 Best Local Similarity 50.3%; Pred. No. 0.00028;
 Matches 151; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

RESULT 10
 ACL29882/c

```

QY 230 GGGGCTTGATGGCTTGGGTTGGTGTGAGAGGCCAACAAATTTTATAAGAAATATATATAAT 289
DB 33712 GTGGATGGGATAGTGGTGTTCGGAGGGGAAACCCAGGAAAGGGGATAACATTTGAAATGT 33771
QY 290 AAAAATAAATAAATTTAAAGTGTGGAGTGAAACGGTGGGCGAGAAATATACACAGAA 349
DB 33772 AATAAAGAAATATTTTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 33831
QY 350 GAGTACTTTTAAATGCGCAACCAAGGACAGATTACAACTTGATTTCTTGGACCTCGAATA 409
DB 33832 AGGAAGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA 33891
QY 410 CGAGATATGCTGTGTAAGAATAAAGAGAGAGTGGAGTGCATTTGAAATGAATGAGAG 469
DB 33892 AGAAGGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 33951
QY 470 CGCAAAATGAGGAGACGAATAAATGAATATATATCAAGAGTGCATTTCCCTATTATT 529
DB 33952 AGAAGAAAGAAAGAAAGAAAGAGTGCATTCATGATCAAGATCAAGTTTACCCCTTCT 34011

RESULT 9
LOCUS I66494 7218 bp DNA linear PAT 28-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION I66494
VERSION I66494.1 GI:2724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner, F., Scheifflinger, P. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
    source
        1..7218
            /organism="unknown"
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match 4.4%; Score 61; DB 6; Length 7218;
Best Local Similarity 2.2%; Pred. No. 0.00041;
Matches 7; Conservative 200; Mismatches 110; Indels 0; Gaps 0;

QY 603 GCAGTGGGAGTTGCAACATGTACCAACAAATCCATTCATCCAAACCTAAATTTATC 662
DB 1055 GGAGTTGGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1114
QY 663 CTCCTCAATTACTTACCTACACTATACCTAGTAATATGCTCCTGCCTGTAACTCCTC 722
DB 1115 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1174
QY 723 CACTGCTGCACAGCTCTTAGTCAATCCATCTGCTTCAAAATAGGCATTTATTTGTTCT 782
DB 1175 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1234
QY 783 TCCCTCCGACTGAAGGCTATCGACCGACCGACCGCTCATCTCTCTCTCTGCGCAAT 842
DB 1235 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1294
QY 843 TTTTCTGCTGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATG 902
DB 1295 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1354
QY 903 TATCTATCTCTCCCT 919
DB 1355 YYYYYYYYYYYYYYYY 1371

RESULT 10
ACL29882/c

```

```

LOCUS              AC129882              147708 bp      DNA      linear      HTG 23-AUG-2002
DEFINITION        Ornithorhynchus anatinus clone CLM1-265D21, WORKING DRAFT SEQUENCE,
4 ordered pieces.
ACCESSION         AC129882
VERSION          AC129882.2  GI:22450573
KEYWORDS         HTG; HTGS PHASE2; HTGS DRAFT
SOURCE           Ornithorhynchus anatinus (platypus)
ORGANISM         Ornithorhynchus anatinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
1 (bases 1 to 147708)
   Akhter N., Antonellis A., Ayele K., Beckstrom-Sternberg S.M.,
   Benjamin B., Blakeley R.W., Bouffard G.G., Breen K., Brinkley C.,
   Brooks S., Dietrich N.L., Granite S., Guan X., Gupta J.,
   Haghighi P., Han J., Hansen N., Ho S.-L., Idol J.R., Karlins E.,
   Laric P., Lee-Lin S.-Q., Legaspi R., Maduro Q.B., Maduro V.B.,
   Margulis E.H., Masello C., Maskeri B., Mastrlian S.D.,
   McCloskey J.C., McDowell J., Pegurigan C., Pearson R.,
   Portnoy M.E., Prasad A., Reddix-Dugue N., Schueler M.G., Sison C.,
   Stantripop S., Thomas J.W., Thomas P.J., Touchman J.W., Vogt J.L.,
   Wetherby K.D., Wiggins L., Young A., Zhang L.-H. and Green E.D.
   NISC Comparative Sequencing Initiative
2 (bases 1 to 147708)
   Green E.D.
   Direct Submission
   Submitted (06-AUG-2002) NIH Intramural Sequencing Center, 8717
   Government Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 147708)
   Green E.D.
   Direct Submission
   Submitted (23-AUG-2002) NIH Intramural Sequencing Center, 8717
   Government Circle, Gaithersburg, MD 20877, USA
   On Aug 23, 2002 this sequence version replaced gi:22122885.
----- Genome Center
   Center: NIH Intramural Sequencing Center
   Center code: NISC
   Web site: http://www.nisc.nih.gov
   Contact: nisc\_zooemhgri.nih.gov
----- Project Information
   Center project name: dwc
   Center clone name: 265D21

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

Summary statistics

Sequencing vector:	plasmid; n/a; 100% of reads
Chemistry:	Dye-terminator Big Dye; 100% of reads
Assembly program:	Phrap; version 0.990319
Consensus quality:	146577 bases at least Q40
Consensus quality:	147063 bases at least Q30
Consensus quality:	147215 bases at least Q20
Insert size:	148000; agarose- ρ p
Insert size:	147408; sum-of-contigs
Quality coverage:	9.31x in Q20 bases; agarose- ρ p
Quality coverage:	9.35x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.

```

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 87687: contig of 87687 bp in length
* 87688 87787: gap of unknown length
* 87788 90934: contig of 3147 bp in length
* 90935 91034: gap of unknown length
* 91035 124474: contig of 33440 bp in length
* 124475 124574: gap of unknown length
* 124575 147708: contig of 23134 bp in length.
FEATURES
    source
        1..147708
            /organism="Ornithorhynchus anatinus"
            /mol_type="genomic DNA"
            /db_xref="taxon:9258"
            /clone="CLM1-265D21"
            /clone_lib="CLM1"
            1..87687
                /note="assembly_fragment
                clone_end:T7
                vector_side:left"
            1..34612
                /note="clone overlaps with GenBank Accession Number
                AC130185 clone CLM1-379P1 (center project name dwt
                74481..147708
                /note="clone overlaps with GenBank Accession Number
                AC127463 clone CLM1-27L6 (center project name dwt
                87768..90934
                /note="assembly_fragment"
                91035..124474
                /note="assembly_fragment"
                124575..147708
                /note="assembly_fragment
                clone_end:Sp6
                vector_side:right"
    40144 a 33940 c 33579 g 39745 t 300 others
BASE COUNT
ORIGIN

```

Query Match	4.1%	Score 57.6;	DB 2;	Length 147708;
Best Local Similarity	50.0%;	Prod. No. 0.0027;		
Matches 144;	Conservative	0;	Mismatches 144;	Indels 0; Gaps 0;
QY	226	TGTCGGGGCTTCGATCGCTTGGGTTTCGGTGTGAGAGGCCAACAAATTTATAAGAAATATATA	285	
Db	88401	TGTCGGGGGTGGGGTATATGAAATGTGTGAGAGAGAGCAAAATATAAGAAAAATA	88342	
QY	236	AAATATAAAAAATAAAAAATTTTAAGTGTGTGGAAGTGAAAAACGGTGGGCGCAAAATATACAC	345	
Db	88341	AAGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAA	88282	
QY	346	AGAAAGAGTACTTTTAAACAATGCCCAACCAAGGCGAGATTCAACAATCTTGATTTCTCGACCTCG	405	
Db	88281	AGAGAGAAAGAGAGAAAGAA	88222	
QY	406	AATACGAGATATCTGGTGTAAAGAAATAAAGAGAGAGTGGAGTGCATTTGAAATGTAATCG	465	
Db	88221	AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAG	88162	
QY	466	AGAGCGCACAAAATGGAGGACGCAATTAATGAAATATATATGCAAGAGTG	513	
Db	88161	AGAAAGAGAGAAAGAGAGAAAGAGAGAAAGAGAGAAAGAGAGAGAGAGAGAGAGAGAG	88114	

RESULT 11
AC074305

LOCUS 305	AC074305	210877 bp	DNA	linear	HTG 26-JUL-2000
DEFINITION	Mus musculus chromosome 16 clone RP23-1012, WORKING DRAFT SEQUENCE.				
	41 unordered pieces.				
ACCESSION	AC074305				
VERSION	AC074305.1	GI:9454521			
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				

AC074303	LOCUS	AC074305	210677 bp	DNA linear	HTG 26-JUL-2000
	DEFINITION	Mus musculus chromosome 16 clone RP23-1012, WORKING DRAFT SEQUENCE,			

ACCESSION	AC074305	41 unordered pieces.
VERSION	AC074305.1	GI:9454521
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

Db 7167 AGAAAGAGGAAA 7213

RESULT 12

AC134879/C

LOCUS

DEFINITION

AC134879

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC134879 199033 bp DNA linear HTG 23-FEB-2003

Homo sapiens chromosome UNK clone RP11-295P22, WORKING DRAFT

SEQUENCE, 13 unordered pieces.

AC134879

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 199033)

The sequence of Homo sapiens clone

Unpublished

2 (bases 1 to 199033)

Waterston, R.H.

Direct Submission

Submitted (01-OCT-2002) Genome Sequencing Center, Washington

University School of Medicine, 444 Forest Park Parkway, St. Louis,

MO 63108, USA

3 (bases 1 to 199033)

Waterston, R.H.

Direct Submission

Submitted (23-FEB-2003) Genome Sequencing Center, Washington

University School of Medicine, 444 Forest Park Parkway, St. Louis,

MO 63108, USA

On Feb 23, 2003 this sequence version replaced gi:23396428.

----- Genome Center -----

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: submissions@wustl.edu

----- Project Information -----

Center project name: H_NH0295P22

----- Summary Statistics -----

Sequencing vector: M13; 0%

Sequencing vector: plasmid; 100%

Chemistry: Dye-primer ET; 0% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 193954 bases at least Q40

Consensus quality: 195124 bases at least Q30

Consensus quality: 195874 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently

* consists of 13 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 2110: contig of 2110 bp in length

2111 2210: gap of unknown length

2211 4639: contig of 2429 bp in length

4640 4739: gap of unknown length

4740 10008: contig of 5269 bp in length

10009 10108: gap of unknown length

10109 13836: contig of 3728 bp in length

13837 13936: gap of unknown length

13937 17705: contig of 3769 bp in length

17706 17805: gap of unknown length

17806 24927: contig of 7122 bp in length

24928 25027: gap of unknown length

25028 31944: contig of 6917 bp in length

31945 32044: gap of unknown length

* 32045 50505: contig of 18461 bp in length

* 50605: gap of unknown length

* 50606 63617: contig of 13012 bp in length

* 63618 63717: gap of unknown length

* 63718 82874: contig of 19157 bp in length

* 82875 82974: gap of unknown length

* 82975 106052: contig of 23078 bp in length

* 106053 106152: gap of unknown length

* 106153 131178: contig of 25026 bp in length

* 131179 131278: gap of unknown length

* 131279 199033: contig of 67755 bp in length.

Location/Qualifiers

1. 199033

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="UNK"

/clone="RP11-295P22"

1. 2110

/note="assembly_name:Contig17"

2211. 4639

/note="assembly_name:Contig18"

4740. 10008

/note="assembly_name:Contig19"

10109. 13836

/note="assembly_name:Contig20"

13937. 17705

/note="assembly_name:Contig21"

17806. 24927

/note="assembly_name:Contig22"

25028. 31944

/note="assembly_name:Contig23"

32045. 50505

/note="assembly_name:Contig24"

50606. 63617

/note="assembly_name:Contig25"

63718. 82874

/note="assembly_name:Contig26"

82975. 106052

/note="assembly_name:Contig27"

106153. 131178

/note="assembly_name:Contig28"

131279. 199033

/note="assembly_name:Contig29"

BASE COUNT 57536 a 42075 c 35735 g 62476 t 1211 others

ORIGIN

Query Match 3.9%; Score 55; DB 2; Length 199033;

Best Local Similarity 58.1%; Pred.No. 0.012;

Matches 97; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 178 AAAGAAGTAAAGGTAATTCAGTCTAGTCAAGTACATTCGATTCGCTTGTGCGGGCTTG 237

Db 154397 AATGGAATCAATGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 154338

QY 238 GATCGCTGGTTCGTTGTGAGAGCCCAATTTATTAAGAAATATATAATAATAATAATA 297

Db 154337 AATGGAATTCGTTGTGAGAGCCCAATTTATTAAGAAATATATAATAATAATAATA 154278

QY 298 AAAAAATTAAGTCTTGAAGTGAAGCGGTGGGGCAGAAATATACA 344

Db 154277 AAGAATTTATGAGTGTGATGGAACATATGGAATGGAATGGAATGGAATGGAATG 154231

RESULT 13

AL928592

LOCUS

DEFINITION

Mouse DNA sequence from clone R24-168J7 on chromosome 4, complete

sequence.

ACCESSION

AL928592

VERSION

AL928592.13 GI:31071444

KEYWORDS

HTG.

SOURCE

Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Giselle, H.
1 (bases 1 to 126611)

REFERENCE 1 (bases 1 to 126611)
AUTHORS Direct Submission
TITLE Submitted (22-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquy@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 23, 2003 this sequence version replaced gi:29125348.
COMMENT Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: UK Medical Research Council
Center code: UK-MRC
Web site: http://mrcseq.har.mrc.ac.uk
Contact: mouse@har.mrc.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30);
an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep/ RP24-168J7 is
from a Male (C57BL/6J) mouse BAC library VECTOR: p7ARBAC1.

FEATURES
source Location/Qualifiers
1..126611
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP24-168J7"
/clone_lib="RPCI-24"
BASE COUNT 40625 a 24197 c 24348 g 37441 t
ORIGIN

Query Match 3.8%; Score 53.8; DB 10; Length 126611;
Best Local Similarity 49.8%; Pred. No. 0.023;
Matches 136; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 230 GGGGCTTGGATGGTTCGGTTCGTGTCAGAGCCACATTTTATAAGAAATATATAAAAT 289
Db 31351 GTGGTTGGGAGCAGGCGAGGAGATATAGAGGTTTATAGGATGCAATTTGAAAT 31410
QY 290 AAAAAATATAAAATTTAAGTTGGTAAGTGAACCGTGGCGGAGAAATATACAGAA 349
Db 31411 GTAATGAAAAATATTATTAATAAAATTTTCAAAAAGAAAGAAAGAAAGAAAGAA 31470
QY 350 GAGTACTTTTCAATGGCCACCAAGCAGATTCAACACTTGATTTCTGACCTCGAATA 409
Db 31471 AGAAGAA 31530
QY 410 CGAGATATGTTGGTAAAGAAATAAAGGAAGTGGATGCGATTTGAAATGAATGGAGAG 469
Db 31531 AGAAGAA 31590
QY 470 CGCACAATCGAGACGAATTAATGAAATATA 502

Db 31591 AGAAGAA 31623

RESULT 14
BX004853
LOCUS
DEFINITION
BX004853 181093 bp DNA linear HTG 03-MAY-2003
Danio rerio clone DKEY-13011, *** SEQUENCING IN PROGRESS ***, 2
uncloned pieces.
ACCESSION
BX004853 GI:30349944
VERSION
HTG; HTGS PHASE1; HTGS ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Danio rerio (zebrafish)
SOURCE
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 181093)
Harrison, E.
Direct Submission
Submitted (02-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On May 3, 2003 this sequence version replaced gi:30024305.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zkl3011
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 180623 bases at least Q40
Consensus quality: 180751 bases at least Q30
Consensus quality: 180906 bases at least Q20
Insert size: 17880; sum-of-contigs
Insert size: 17280; 4.9% error; agarose-fp
Quality coverage: 6.77x in Q20 bases; sum-of-contigs Quality
coverage: 7.52x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 95025: contig of 95025 bp in length
* 95026 95125: gap of 100 bp
* 95126 181093: contig of 85968 bp in length.
* Location/Qualifiers
1..181093
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-13011"
/clone_lib="DanioKey"
1..95025
/note="assembly_fragment:00925"
95126..181093
/note="assembly_fragment:02465.0"
BASE COUNT 56902 a 34623 c 33889 g 55579 t 100 others
ORIGIN

Query Match 3.8%; Score 53.2; DB 2; Length 181093;
Best Local Similarity 51.2%; Pred. No. 0.033;
Matches 124; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 257 AGAGCCACCAATTTTATAGAAATATATAAAATATAAAATATAAAATTTAAGTGTGGA 316
Db 81583 AAAAGGAATCTAAGCAAAATCAATGGAACAAAGAAATATAAAATTTAACTGAA 81642

QY 317 AGTGAACCGTGGGCGAATATATACACAGAGAGTAGTACTTTAAACAATCGCGCAACCAAGG 376
 Db 81643 ATATGACACAGACAG 81702
 QY 377 CAGATTCACACTTGATTTCTTGACCTCGAATACAGAGATAATGGTGGTGAAGAAATAAAGG 436
 Db 81703 CAGACAGG 81762
 QY 437 AGAGTGGAGTGCATTGTAATGAATGAATGAGAGCGCACAAATATGGAGCAGCAATATGA 496
 Db 81763 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 81822
 QY 497 AA 498
 Db 81823 AA 81824
 RESULT 15
 AC111508/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-133F12, WORKING DRAFT SEQUENCE.
 AC111508
 ACCESSION
 VERSION AC111508.5 GI:30579329
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 210483)
 AUTHORS Muzny, D., Marie, H., Metzker, M., Jee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Altschul, S., Amin, A., Anguiano, D., Anylebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hernandez, K., Georgergias, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, M., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebowitz, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhera, L., Louised, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mathew, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nait, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankwelu, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puzo, M., Quirio, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G., and Gibbs, R.A.
 Unpublished
 Direct Submission
 2 (bases 1 to 210483)
 Direct Submission
 Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 210483)
 Direct Submission
 Submitted (13-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 On May 13, 2003 this sequence version replaced gi:24818983.
 The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GMLR
 Center clone name: CH230-133F12
 ----- Summary Statistics
 Assembly program: Atlas 3.0;
 Consensus quality: 202700 bases at least Q40
 Consensus quality: 203861 bases at least Q30
 Consensus quality: 204681 bases at least Q20
 Estimated insert size: 211532; sum-of-contigs estimation
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length.
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * the accession number will be 210483 bp in length.

 1. 210483: contig of 210483 bp in length.
 Location/Qualifiers
 1. 210483
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-133F12"
 2339..3169
 /note="clone_boundary
 clone_end:17
 site:ECORI
 end_sequence:BH324644"

 misc_feature
 source
 FEATURES
 source
 misc_feature

Thu Jan 22 09:08:29 2004

BASE COUNT	52766 a	43717 c	45649 g	63169 t	5182 others
ORIGIN					

Query Match 3.8%; Score 53; DB 2; Length 210483;

Best Local Similarity 52.0%; pred. No. 0.037;

```

BEST LOCAL SIMILARITY 32.0%; 110; Indels 0; Gaps 0;
Matches 119; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

```

SAUCYBPM /ETT COMPTASTOC /ATT CORRUMITT /ATT /ATT

274 AAGAAATATATAAAATAAAAAATTTAAGTGTTCGAAAGTGAAACGGTGGGC 333

[illegible]

334 AGAAATATACACAGAAGAGTACTTTAACAAATGCGCAACCAAGGCAGATTCACTTGAT 393

334 AGAATATATACACAGAAGAGTACTTTTACCAATGCGCACCCAGAGCAGATTCACACACTTGT 335

204213 AGAAAGAAAGAAAGAGGGAAGAAAGAGAAAGAAAGAAAGAA 204154

[illegible]

394 TTCTGGACCTCGAATACGAGATAATGCTGTTAAGAAATAAAGGAAGAGTGGAGTGCATTT 453

394 TTC TGG ACC TCG AAT AC GAG ATA ATGG TGG TAA GAA ATA TAG GAA GAG TGG AGA TGG CAT TTT TTT

Pb 204153 AGAAGGAGAAGAAAGGAAGGAAGGAGGAAGGAA 204094
DH

[illegible]

454 CCAAAATGAAATCGAGAGCGGCACAAAATCGAGGACGAATAAATGAAATATA 502

454 GAAAATGAAATGGAGAGCGCACAAAATGGAGGACGAAATAAATGHAATAATA 302
QY

204093 GGAAAGAAACGAGAGAGAGAAAGAAAGAAAGAAAGAAAGAAA 204045

Search completed: January 21, 2004, 22:14:19
Job time : 5113.33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 16:52:57 ; Search time 61.4877 Seconds
(without alignments)
6525.161 Million cell updates/sec

Title: US-09-936-869-3

Perfect score: 909

Sequence: 1 atgggtcgtggaagattga.....cgcattgtattcttctgttg 909

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1:39956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	260	28.6	1219	4	US-09-410-464-10
2	239.8	28.6	723	4	US-09-410-464-11
3	255.6	28.1	959	3	US-09-067-800-7
4	255.6	28.1	959	3	US-09-349-677-7
5	254.6	28.0	1125	1	US-08-460-512-6
6	248.2	27.3	714	4	US-09-410-464-15
7	248.2	27.3	1159	4	US-09-410-464-14
8	248.2	27.3	1297	4	US-09-433-579-1
9	245	27.0	896	3	US-09-087-800-5
10	245	27.0	896	3	US-09-349-677-5
11	232.4	25.6	1457	1	US-08-460-512-1
12	230.8	25.4	1097	1	US-08-460-512-3
13	147.2	16.2	1043	2	US-08-867-087B-12
14	141.6	15.6	1062	3	US-09-067-800-1
15	141.6	15.6	1062	3	US-09-105-652-1
16	141.6	15.6	1062	3	US-09-349-677-1
17	141.6	15.6	1062	4	US-08-788-564-1
18	138	15.2	1345	1	US-08-592-214A-7
19	138	15.2	1345	3	US-08-659-188-7
20	138	15.2	1345	3	US-08-655-227-7
21	138	15.2	1345	3	US-08-655-241-7
22	138	15.2	1345	3	US-09-149-976-7
23	138	15.2	1345	3	US-09-398-326-7
24	131.2	14.4	795	3	US-08-904-284-2
25	129.2	14.2	1180	2	US-08-867-087B-16
26	126.8	13.9	794	1	US-08-592-214A-3
27	126.8	13.9	794	3	US-08-659-188-3

ALIGNMENTS

RESULT 1.
US-09-410-464-10
; Sequence 10, Application US/094.0464
; Patent No. 6395892
; GENERAL INFORMATION:
; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; FILE REFERENCE: 53375
; CURRENT APPLICATION NUMBER: US/09/410.464
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: 09/287,700
; EARLIER FILING DATE: 1993-04-06
; EARLIER APPLICATION NUMBER: 60/080,851
; EARLIER FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 10
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Populus balsamifera subsp. trichocarpa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (196)..(921)
US-09-410-464-10

Query Match	28.6%	Score	260;	DB	4;	Length	1219;
Best Local Similarity	63.7%	Pred. No.	5.5e-67;				
Matches	433;	Conservative	0;	Mismatches	235;	Indels	12;
Gaps	2;						
QY	2	TGGTCTGGGAGATTGAGATAAGAGGATTGAAATCTAGAACCGACGCTCATT	61				
DB	242	TGGGAGCGGAAAGGTGGAGATCAAGCGGATCGAGAACCAACCAACCAAGTCACTT	301				
QY	62	TCTCAAGCGCGGAAATGGTTTATTAAAGAGCGGTATGAATTCAGTTCTTTTGTGATG	121				
DB	302	TCTCAAGAGCGGAGTGGTTTCTCAAGAAAGCTTACGAATATCTGTTCTTTGGGATG	361				
QY	122	CAGAAGTGGCGCTCATGCTCTTCCAGAGAGGAGACTTTATGAATTTGCCAACCA	181				
DB	362	CTGAGGTGCACTCATGCTCTTCTAGCGCGGTTCGCTTTATGAGTACTCTTAACGATA	421				
QY	182	CGCTGAAGAGAGGAGATTGAGAGGTACAAGAGACTTGGCGTTGACACCAACCGAGGGG	241				
DB	422	GTGTCAATCAACAAATTTGAGAGGTACAAAAGGCACTTCGAGATTCTTCAACACTGGGT	481				
QY	242	CGATATCAGAGTCCAATTTCTCAGTATGCGCAAGAGGCTGGTAACTCAGACACACA	301				
DB	482	CTGTTTCTGAAGCAATGCTCAGTACTACGAGCAAGAGTCCCAAGCTGGGTTCCTCCAAA	541				


```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/349,677
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/067,800
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 959 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 78..818
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..959
; OTHER INFORMATION: /note= "AGL5 cDNA and deduced
; OTHER INFORMATION: protein sequences."
; US-09-349-677-7

Query Match      28.1%; Score 255.6; DB 3; Length 959;
Best Local Similarity 68.3%; Pred. No. 9.9e-66;
Matches 354; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY   1 ATGGGTCGTGGGAAGATTGAGATAAAGAGAGATTGAAATTAACCTACGAAACCGACAGGCCTACT 60
DB   123 ATAGGGAGAGGAAGATAGAGATAAAGAGAGATAGAGAACAACACAGCAATCGTCAAGTCAC 182
QY   61 TTCTGCAAGCGCCGAATGGTTTATTAAAGAAAGGCGTATGAATATCAGTCTCTTGTTGAT 120
DB   183 TTCTGCAAAACGACGAATGGTTTACTCAAGAAAGCTTATGAGCTCTCTCTCTGTTGATGAC 242
QY   121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGAGACTTTATGAATTCGCCAACAC 180
DB   243 GCTGAGGTTGCTCTTGTGTCATCTTCTCCACTCGAGGCGCTCTCTACGAGTAGCCCAACAC 302
QY   181 AGCGTGAAGAGGACGATTGAGAGTACAGAGACATTCGTTGACAAACACCAGCGAGG 240
DB   303 AGTGTGAGAGGAACAATAGAAAGAGGTACAAAGAAAGCTTGTCCGACGCCCTTAACCTCCG 362
QY   241 GCGATATCAGAGTCCCAATTCAGTATTGGCAACAGAGGCGTGGTAAATCTCAGAACACAG 300
DB   363 ACCATCACCGAAGCTAATCTAGTACTATCAGCAAGAGGCGTCTAAATCTCCGGAGACAG 422
QY   301 ATTGACATTTGCAAAATGCAATAGGCAATTCATGGGTGACGGGCTTACAGCTTTGAC 360
DB   423 ATTCGGGACATTCAGAAATTTGAACAGACATTCITGGGTGAATCTCTTGGTTCCTTGAAC 482
QY   361 ATTAAGGAACCTCAAGCAACTTGGAGTTTCGACTTTGAAAAGAGGAATCAGCCGAGTGCATCC 420
DB   483 TTTAAAGGAACCTCAAGAACCTTGAAGTAGGCTTTGAGAAAGGAATCAGTCTGTCGATCC 542
QY   421 AAAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATCGAGAGGAGGAACACATAC 480
DB   543 AAGAGCAGCAGATGTTAGTTGAGAGATTTGAATATCATGCAAAAAAGGAATTCAGCTG 602
QY   481 ATCCAGGAGAAATGAGATCTTCGACGAGCAAGATAGCCGA 518
DB   603 CAAAACGATAACATGATCTCCGCTCCAGATTAACGA 640

RESULT 4
US-09-349-677-7
; Sequence 7, Application US/09349677
; Patent No. 6288305
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin P.
; APPLICANT: Ferrandis, Cristina
; TITLE OF INVENTION: Seed Plants Characterized by Delayed
; TITLE OF INVENTION: Seed Dispersal
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```



```
RESULT 5
US-08-460-512-6
; Sequence 6, Application US/08460512
; Patent No. 5744693
; GENERAL INFORMATION:
; APPLICANT: MEYEROWITZ, Elliot M.
; APPLICANT: YANOFKY, Martin F.
; APPLICANT: WA, Hong
; TITLE OF INVENTION: PLANTS HAVING ALTERED FLORAL DEVELOPMENT
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard F. Trecartin
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,512
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/293,278
; FILING DATE:
; APPLICATION NUMBER: US/07/956,694
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-57322/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1125 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 110..853
; US-08-460-512-6

Query Match 28.0%; Score 254.6; DB 1; Length 1125;
Best Local Similarity 67.8%; Pred. No. 2.1e-65;
Matches 356; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 2 TGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAATATCTAGCAACCGACAGGTCACCTT 61
DB 159 TGGGAGAGGAAGATTGAGATCAACGGATCGAAACACAAAGATCGTCAAGTCACCTT 218
QY 62 TCTGCAAGCCGCGAAATGGTTTATAAAGAGCGGTATGAATATCAGTTCTTTGTGATG 121
DB 219 TCTGCAAGAGACGGAATGGTTTACTCAAAAAGGCGCTATGAATATCTGTGCTCTGTGATG 278
QY 122 CAGAAAGTGGCCCTCATGCTCTTCTCCAGCAGAGGAGACCTTATGAATTTGCCAACCA 181
DB 279 CTGAGTTGCTTTGATTGCTCTCTCAAGCAGAGGACAGCTCTATGATGATGCCAACCA 338
QY 182 GCGTGAAGAGGACGATTGAGAGGTACAAAGACCTTTCGGTTGACCAACACCGAGGGG 241
DB 339 CTGTGAAGACCAATTTGAGAGGTACAAAGAGCTTGTTCAGATTCCTCAAAACACTGGTT 398
QY 242 CGATATCAGACTCCAAATTTCTCAGTATTGGCAACAGAGGCGTGTAAACTCAGACACAGA 301

; Query Match 27.3%; Score 248.2; DB 4; Length 714;
; Best Local Similarity 67.0%; Pred. No. 1.3e-63;
; Matches 352; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 2 TGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAATATCTAGCAACCGACAGGTCACCTT 61
DB 47 TGGGAGGGGAAAGGTGGAGATCAACGGATCGAGAACACCACCAATGCCAAGTCACCTT 106
QY 62 TCTGCAAGCCGCGAAATGGTTTATAAAGAGCGGTATGAATATCAGTTCTTTGTGATG 121
DB 107 TCTGCAAAAGGCGGAATGGTTTGTCTCAAGAAAGCCTATGAATATCTGTCTTTGCGATG 166
QY 122 CAGAAAGTGGCCCTCATGCTCTTCTCCAGCAGAGGAGACCTTATGAATTTGCCAACCA 181
DB 167 CTGAGGTTGCACTCATGCTCTTCTCCAGCGGTGGACGCCCTTATGATGATCTTAACAATA 226
QY 182 GCGTGAAGAGGACGATTGAGAGGTACAAAGAGCTTTCGTTGACCAACACCGAGGGG 241
DB 227 GTGTCAAAATCTCAATTTGAAAGGTACAAAAGGCGATGTGCAGATTCTTCCAAACAGGGT 286
QY 242 CGATATCAGAGTCCAAATTTCTCAGTATTGGCAACAGGAGGCTGTAACTCAGACACAGA 301
DB 287 CAGTTCTTGAAGCCAAATGCTCAGTTCTTATCAGCAAGAGCTGCCAGCTGCCCTCGCAAA 346
QY 302 TTGACATTTTGCAAAATGCAAAATFAGGCAATTTGATGGGTGACGGGCTTACAGCTTTGAACA 361
DB 347 TTGCTAATTTGCAGAAATTCAAAACAGGAAATATGCTGGGTGAATCACTTAGTGCATTGAGTG 406

RESULT 6
US-09-410-464-15
; Sequence 15, Application US/09410464
; Patent No. 6395892
; GENERAL INFORMATION:
; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; FILE OF INVENTION: poplar and other plant species.
; FILE REFERENCE: 53375
; CURRENT APPLICATION NUMBER: US/09/410,464
; CURRENT FILING DATE: 1999-10-01
; EARLIER APPLICATION NUMBER: 09/287,700
; EARLIER FILING DATE: 1999-04-06
; EARLIER APPLICATION NUMBER: 60/080,851
; EARLIER FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 714
; TYPE: DNA
; ORGANISM: Populus balsamifera subsp. trichocarpa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(714)
; US-09-410-464-15

Query Match 27.3%; Score 248.2; DB 4; Length 714;
Best Local Similarity 67.0%; Pred. No. 1.3e-63;
Matches 352; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 2 TGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAATATCTAGCAACCGACAGGTCACCTT 61
DB 47 TGGGAGGGGAAAGGTGGAGATCAACGGATCGAGAACACCACCAATGCCAAGTCACCTT 106
QY 62 TCTGCAAGCCGCGAAATGGTTTATAAAGAGCGGTATGAATATCAGTTCTTTGTGATG 121
DB 107 TCTGCAAAAGGCGGAATGGTTTGTCTCAAGAAAGCCTATGAATATCTGTCTTTGCGATG 166
QY 122 CAGAAAGTGGCCCTCATGCTCTTCTCCAGCAGAGGAGACCTTATGAATTTGCCAACCA 181
DB 167 CTGAGGTTGCACTCATGCTCTTCTCCAGCGGTGGACGCCCTTATGATGATCTTAACAATA 226
QY 182 GCGTGAAGAGGACGATTGAGAGGTACAAAGAGCTTTCGTTGACCAACACCGAGGGG 241
DB 227 GTGTCAAAATCTCAATTTGAAAGGTACAAAAGGCGATGTGCAGATTCTTCCAAACAGGGT 286
QY 242 CGATATCAGAGTCCAAATTTCTCAGTATTGGCAACAGGAGGCTGTAACTCAGACACAGA 301
DB 287 CAGTTCTTGAAGCCAAATGCTCAGTTCTTATCAGCAAGAGCTGCCAGCTGCCCTCGCAAA 346
QY 302 TTGACATTTTGCAAAATGCAAAATFAGGCAATTTGATGGGTGACGGGCTTACAGCTTTGAACA 361
DB 347 TTGCTAATTTGCAGAAATTCAAAACAGGAAATATGCTGGGTGAATCACTTAGTGCATTGAGTG 406
```

QY 362 TTAAGGAACCTCAAGCAACTTGAGGTTTCGACTTTGAAAGGAATCAGCCGAGTGGGATCCA 421
 Db 407 TGAAGGAACCTTAAGAGCTTGGAGATAAACTTGAAGGAATGTTAGTAATTCGTTGGA 466
 QY 422 AAAAGAACAGATGTTGCTTGAAGATCGACATCATGACAGAGAGGGAACACATACCTTA 481
 Db 467 AAAAGAAATGAGCTGTTGTTGCTGAAATGAGTATATGAGAGAGGAGATTGACTTGC 526
 QY 482 TCAGAGAGAAATGAGATTCTTCAGCAAGATAGCCGAGTGTGAGA 526
 Db 527 ACAACATACAGAGCTTCTCCGAGCAAGATTGACAGAGATGAAA 571

RESULT 7
 US-09-410-464-14
 ; Sequence 14, Application US/09410464
 ; Patent No. 6395892
 ; GENERAL INFORMATION:
 ; APPLICANT: Strauss et al.
 ; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
 ; ; TITLE OF INVENTION: Poplar and other plant species.
 ; FILE REFERENCE: 53375
 ; CURRENT APPLICATION NUMBER: US/09/410,464
 ; CURRENT FILING DATE: 1999-10-01
 ; EARLIER APPLICATION NUMBER: 09/287,700
 ; EARLIER FILING DATE: 1999-04-06
 ; EARLIER APPLICATION NUMBER: 60/080,851
 ; EARLIER FILING DATE: 1998-04-06
 ; NUMBER OF SEQ ID NOS: 24
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 14
 ; LENGTH: 1159
 ; TYPE: DNA
 ; ORGANISM: Populus balsamifera subsp. trichocarpa
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (99)..(815)
 US-09-410-464-14

Query Match 27.3%; Score 248.2; DB 4; Length 1159;
 Best Local Similarity 67.0%; Pred. No. 1.6e-63;
 Matches 352; Conservative 0; Mismatches 173; Indels 0; Gaps 0;
 QY 2 TGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAATATCTAGCAACCGACAGGTCACCTT 61
 Db 145 TGGGAGGGAAGGTTGAGATCAAGCGGATCGAAGACCAACCAATCGCCAGTCACTT 204
 QY 62 TCTGCAAGCCCGAAATGTTTATTAAGAGCGGTATGAAATTATCAGTTCTTTGTGATG 121
 Db 205 TCTGCAAAAGGCGGAATGTTGCTCAAGAAAGCCCTATGAAATATCTGTTCTTTGCGATG 264
 QY 122 CAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAAATTTGCCAACACA 181
 Db 265 CTGAGTTGCACTCAICGCTCTTCTCCAGCCGTGAGCGCTTTATGAGTACTCTTAACAATA 324
 QY 182 CGGTGAGAGAGCAGATTGAGAGGTACAGAGACTTTCGTTGTGACAAACACCGAGGGG 241
 Db 325 GTGTCAAAATCTACAAATTTGAAAGGTACAAAAGGCGTGTGAGATTTCTTCAACACCGGT 384
 QY 242 CGATATCAGAGTCCAAATTTCTCAGTATTGGCAACAGAGGCTGGTAAACTCAGACAAACAGA 301
 Db 385 CAGTTTCTGAGGCCAATGCTCAGTTCTATCAGCAGAGGCTGCGAGCTCGCTCCGAAA 444
 QY 302 TTGACATTTTGCAGAAATGCAATAGGCAATTTGATGGGTGACGGGCTTACAGCTTTGAAACA 361
 Db 445 TTGATTAATTTGCAGAAATTTCAACAGGAATATGCTGGGTGAATCATTAGTGTGATGAGTG 504
 QY 362 TTAAGGAACCTCAAGCAACTTGAGGTTTCGACTTTGAAAGGAATCAGCCGAGTGGGATCCA 421
 Db 505 TGAAGGAACCTTAAGAGCTTGGAGATAAACTTTGAAAGGAATTTGGTAGAATTCGTTGGA 564
 QY 422 AAAAGAACAGATGTTGCTTGAAGATCGACATCATGACAGAGAGGGAACACATACCTTA 481

Db 565 AAAAGAAATGAGCTGTTGTTGCTGAAATTCAGTATATGACAGAGAGGAGATTGACTTGC 624
 QY 482 TCAGAGAGAAATGAGATTCTTCGAGCAAGATAGCCGAGTGTGAGA 526
 Db 625 ACAACATAACCAAGCTTCTCCGAGCAAGATTGACAGAGATGAAA 669
 RESULT 8
 US-09-433-579-1
 ; Sequence 1, Application US/09433579
 ; Patent No. 6444877
 ; GENERAL INFORMATION:
 ; APPLICANT: Rottmann, William H.
 ; TITLE OF INVENTION: LSAG Gene
 ; FILE REFERENCE: LSAG Gene
 ; CURRENT APPLICATION NUMBER: US/09/433,579
 ; CURRENT FILING DATE: 1999-11-04
 ; NUMBER OF SEQ ID NOS: 36
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1297
 ; TYPE: DNA
 ; ORGANISM: Liquidambar styraciflua
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (331)..(1008)
 US-09-433-579-1
 Query Match 27.3%; Score 248.2; DB 4; Length 1297;
 Best Local Similarity 62.6%; Pred. No. 1.7e-63;
 Matches 426; Conservative 0; Mismatches 243; Indels 12; Gaps 2;
 QY 1 ATGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAATATCTAGCAACCGACAGGTCACCT 60
 Db 331 ATGGGTAGGGGAAGATCGAGATCAAGCGGATCGAAGAACACGACAAACCGCCAGTCAAC 390
 QY 61 TTCTGCAAGCCCGAAATGTTTATTAAGAGCGGTATGAAATATCAGTTCTTTGTGAT 120
 Db 391 TTCTGTAACCGCGCAACGGCTTACTCAAGAGCCCTATGAAATATCTGTTCTCTGAT 450
 QY 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAAATTTGCCAACAC 180
 Db 451 GCCGAGGTAGCCCTTATCGTCTTCTACCGTGGCGCTTACGAGTATGCCAACAAAC 510
 QY 181 AGCGTGAAGAGGACGATTGAGAGGTACAGAGACTTTCGTTGTGACAAACACCGAGGG 240
 Db 511 AGTGTAAAGCAACAATTGAGAGGTATAGAGGAGCTCTGTGGATTCTCTCAATACTGGA 570
 QY 241 CGGATATCAGAGTCCAAATTTCTCAGTATTGGCAACAGAGGCTGGTAAACTCAGACAAAC 300
 Db 571 TCTGTTTCTGAGCCATGCTCAGTTCTACAGCAAGAGCTGCAAACTCGGTAACCAA 630
 QY 301 ATTGACATTTTGCAGAAATGCAATAGGCAATTTGATGGGTGACGGCTTACAGCTTTGAC 360
 Db 631 ATCAGAAATATGCAAGATACAAACAGGACTATGCTGGGTGAGTCTTTGGGCTCTCTGAT 690
 QY 361 ATTAGGAACTCAGCAACTTTCGAGTTCGACTTCAAAAAGGAATCAGCCGAGTGCATCC 420
 Db 691 CCCAAGAACTCAGGGTCTCGAGACTAAATTAGAGAAAGCATTTAGCAAAATAGGTCC 750
 QY 421 AAAAGAACAGAGATGTTGCTTTGAAAGAGATCGACATCATGACAGAGAGGGAACATACCT 480
 Db 751 AAAAGAAATGAGTACTGTTTTCGAAATTTGAATACATGCAAAAAGAGGGAATCGACTTG 810
 QY 481 ATCCAGAGAGATGAGATTCTTCGAGCAGAGATAGCCGAGTGTGAGA---ATAGCCACAAC 537
 Db 811 CACAATGATAACCGATATCTCCGAGCAAAAATAGCTGAAACAGAGAGCTCAACAGCAA 870
 QY 538 ACGAACAATGTTTATCAGCTCCGAAATATGATG-----CACTGCCGCAATTCGACTCT 588
 Db 871 ATGAACTTGTGCGCAGGTGGATCAAACTATGAGGTCTTGGCCATCAGGCCATTTGACTCT 930
 QY 589 CGAAATTTCTACATGCAAAATCTAAATCGATGCGGCCCATCATCTATGCAATCAGGAACAA 648

Db 931 CGGAAGTCTTCCAGTAATGCTTGCACCCCAATCATCACTCTCCCAAGATCAA 990
Qy 649 ACAACGCTTCAGCTGGCTGA 669
Db 991 ATGGCCCTTCAATTAGITTA 1011

RESULT 9

US-09-067-800-5
; Sequence 5, Application US/09067800
; Patent No. 6198024
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Ferrandiz, Cristina
; TITLE OF INVENTION: Seed Plants Characterized by Delayed
; TITLE OF INVENTION: Seed Dispersal
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESS: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/067,800
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 896 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..753
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 896
; OTHER INFORMATION: /note= "There is a poly(A) tail at
; OTHER INFORMATION: the end of the cDNA sequence."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..896
; OTHER INFORMATION: /note= "AGU1 cDNA and deduced
; OTHER INFORMATION: protein sequences."
US-09-067-800-5

Query Match 27.0%; Score 245; DB 3; Length 896;
Best Local Similarity 67.1%; Pred. No. 1.3e-62;
Matches 347; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGGTCTGGGAAGATTGAGATAAGAGGATTGAAATACAGAACCGACAGGTCACCTT 61
Db 53 TAGGAGAGGGAATAAGATAGATTAAGAGGATAGAGACACACAATCTCTCAAGTACTT 112
Qy 62 TCTGCACGCCCGGAAATGGTTTATTAAAGAGCGGTATGAATTATCAGTCTCTTGATG 121

Db 113 TCTCAAAACGACGCAATGGTCTTCTCAAGAAAGCTTATGAACCTCTCTCTGTGTGATG 172
Qy 122 CAGAAGTGGCCCTCATCGTCTCTCCAGCAGAGGGAGCTTTATGAATTTGCCAACCA 181
Db 173 CCGAAGTTGCCCTCGTCACTCTCTCCAGTCTCTCTCTGCGCCGTCTCTATGAGTACGCCAACCA 232
Qy 182 GCGTGAAGAGGACATTTGAGAGGTACAAGAGACTTTCGCTTGACAAACCAACCGAGGG 241
Db 233 GTGTGAGGGGTACAATTTGAAGGTACAAGAAAGCTTGTTCGATGCGCTCAACCCCTCTT 292
Qy 242 CGATATCAGAGTCCAATTTCTCAGTATTGGCAACAGCAGGCTGGTAAACTCAGACACACA 301
Db 293 CCGTCACCGAAGCTTACTCAGTATCTATCAGCAAGAGCCTCTAAGCTTCGAGGCGCA 352
Qy 302 TTGACATTTTGAAGATGCAATAGCATTGATGGGTGACCGGCTTACAGCTTTGAACA 361
Db 353 TTCAGATATTCAAAATTCAAATAGGCATATTGTTGGGAATCAGCTTGGTTCCTTGA 412
Qy 362 TTAAGGAACCTCAAGCAACTTGAAGTTTGGACTTGAAGAAAGGAATCAGCCGAGTCCGATCCA 421
Db 413 TCAAGGAACCTCAAAACCTAGAGGAGCTTCTGAAAAGAGATCAGCCGCTTCCGCTCCA 472
Qy 422 AAAAGACGAGATGTTGCTTTGAAGAGATCGACATCATGACAGAGAGGGGAACACACTTA 481
Db 473 AAAAGATGAGCTGTTAGTGGCAGACATAGATATATGACAGAGAGGGGAATGAGTTGC 532
Qy 482 TCCAGGAGAAATGAGATTCTTTCGCGACGCAAGATAGCCGA 518
Db 533 AACACAATAACATGTACTCTCGGAGCAAGATAGCCGA 569

RESULT 10

US-09-349-677-5
; Sequence 5, Application US/09349677
; Patent No. 6288305
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Ferrandiz, Cristina
; TITLE OF INVENTION: Seed Plants Characterized by Delayed
; TITLE OF INVENTION: Seed Dispersal
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESS: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/349,677
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/067,800
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 896 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

RESULT 12
US-08-460-512-3
Sequence 3, Application US/08460512
Patent No. 5744693
GENERAL INFORMATION:
APPLICANT: MEYEROWITZ, Elliot M.
APPLICANT: YANOFKY, Martin F.
APPLICANT: MA, Hong
TITLE OF INVENTION: PLANTS HAVING ALTERED FLORAL DEVELOPMENT
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard F. Trecartin
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,512
FILING DATE: 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/293,278
FILING DATE: 1997
APPLICATION NUMBER: US/07/956,694
FILING DATE: 1997
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-57322/RFT
TELEPHONE: (415) 791-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1097 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 126..884
US-08-460-512-3

Query Match 25.4%; Score 230.8; DB 1; Length 1097;
Best Local Similarity 64.7%; Pred. No. 2.1e-58;
Matches 343; Conservative 0; Mismatches 187; Indels 0; Gaps 0;
QY 4 GGTCTGGGAGATTGAGATAAGAGAGATTGAAATACTACGAAACCGACAGGTCATTTC 63
DB 177 GGGAGAGGAAGATCGAAATAAAACCGATCGAGAACACACGAAACCGTCAAGTTACTTTC 236
QY 64 TGAAGGCGGAAATGTTTATTAAGAGCGGTATGAATATCAGTTCTTTCTGATGCA 123
DB 237 TGAAGGCGGAAATGTTTCTCAAGAAAGCTTACGAATCTCTGTTCTTTGATGCT 296
QY 124 GAACTGGCCCTCATCTGCTTCTCCAGCAGAGGAGACTTTATGAATTCGCAACACAGC 183
DB 297 GAACTGGCACTCATCTGCTTCTCTAGCGTGGCGTCTCTATGATGACTCAACACAGC 356
QY 184 GTGAAGAGGAGATTGAGAGGTACAGAGACTTGGCTTGACACACACGAGGGGGCG 243
DB 357 GTAAAAGGAGCAATTTGAGAGGTACAGAAAGCAATATCGGATAATTTTAAACCGGATCC 416
QY 244 ATATCAGAGTCCAAATTTCTCAGTATTGGCAACAGGAGGCTGGTAAACTCAGACACAGATT 303

DB 417 GTGGCAGAAATTAATGCACAGTATTATCAACAGAAATCTGCCAAATTCGTCACAAATT 476
QY 304 GACATTTTGCAAAATGCAATAGGCAATTTGATGGGTGACGGCTTACAGCTTTGAACATT 363
DB 477 ATCAGCATACAGAACTCGAACAGGCAATTTGATGGGTGAGACGATTTGGTCAATGTCTCC 536
QY 364 AAGGAATCAAGCAACTTGAAGTTCGACTTGAAGAGGAATCAGCCGAGTGGGATCCAAA 423
DB 537 AAGAGCTCAGGAACCTTGAAGAGGAGATTAGACAGAGTGTTAATCGAATCCGATCCAG 596
QY 424 AAGAACGAGATGTTGCTTGAAGAGATCGACATCATCGACAGAGAGGGAACACATATTATC 483
DB 597 AAGAACGACTCTTATTCGCCGAATTTACTACATCGACAGAGAGAGTGTGATTTGTCAT 656
QY 484 CAGAGATGAGATTCTTCCGACAGATAGCCGAGTGTGAGATAGCCA 533
DB 657 AAGGATAACCAAGCTTCTTCTGCTAAGATAGCTGAAATGAGAGGAACAA 706

RESULT 13
US-08-867-087B-12
Sequence 12, Application US/08867087B
Patent No. 5990386
GENERAL INFORMATION:
APPLICANT: An, Gynheung
TITLE OF INVENTION: GENES CONTROLLING FLORAL DEVELOPMENT
TITLE OF INVENTION: AND APICAL DOMINANCE IN PLANTS
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klarquist Sparkman Campbell Leigh &
ADDRESS: Whinston, LLP
STREET: One World Trade Center
STREET: 121 S.W. Salmon Street
STREET: Suite 1600
CITY: Portland
STATE: Oregon
COUNTRY: United States of America
ZIP: 97204
COMPUTER READABLE FORM:
MEDIUM TYPE: Disk, 3-1/2 inch
COMPUTER: IBM PC compatible
OPERATING SYSTEM: MS DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,087B
FILING DATE: June 2, 1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. 08/323,449
FILING DATE: October 14, 1994
APPLICATION NUMBER: U.S. 08/485,981
FILING DATE: June 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Alan E.
REGISTRATION NUMBER: 35,123
REFERENCE/DOCKET NUMBER: 4630-47071
TELEPHONE: (503) 226-7391
TELEFAX: (503) 228-9446
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1043 base pairs
TYPE: nucleic acid
STRANDEDNESS: double stranded
TOPOLOGY: linear
US-08-867-087B-12

Query Match 16.2%; Score 147.2; DB 2; Length 1043;
Best Local Similarity 58.5%; Pred. No. 9.2e-34;
Matches 276; Conservative 0; Mismatches 193; Indels 3; Gaps 1;
QY 1 ATGGGTCTGGGAGATTGAGATAAGAGGATTGAAATACTACGAAACCGACAGGTCCT 60

Db 34 ATGGGAGGGAAGATTGAGCTGAAGCGCATCGAAGCAAGATCAACAGGAGGTACCC 93
Qy 61 TTCTGCAAGCGCGAAATGGTTTATTAAAGAGGGGTATGAATATATATCTTTGTGAT 120
Db 94 TTCTCAAGCGCGCAACGGCTCTCAAGAGGCTACAGAGCTGTCCGTTCTCTGCGAC 153
Qy 121 GGAAAGTGGCCCTCATGCTCTTCTCCAGCAGAGGAGACTTTATGAATTTCCCAACAC 180
Db 154 GCGGAGTGGCGCTCATCTCTTCTCCAGCGCGCAAGCTCTACGAGTTGGCAGCGCC 213
Qy 181 AGCGTCAAGAGGAGCAATTCAGAGGTACAAAGAGACTTGGTTGACAAACCAACCCAGGAGG 240
Db 214 GGCATCAACAGACTTTAGAGGTACCAACATTTGTC---TACAATGCTCAAGATTCC 270
Qy 241 GCGATATCAGTCCAAATCTCAGTATTTGGCAAGAGGCTGTGTAACATCAGACACAG 300
Db 271 AACAATGCACTTTCTGAAACTCAGAGTTGGTTGATACCAATGTCAAAGTTGAAGCAAAA 330
Qy 301 ATTGACATTTGCAAAATCAATAGCATTTGATGGGTGACGGGCTTACAGCTTTGAAC 360
Db 331 TTGAAGCTTTGAGCGCACTCAAGGCACTTGTCTTGGGGAGATCTTGGACCACTCAGC 390
Qy 361 ATTAAGGAATCAAGCAACTTGAGTTTCGACTTTGAAAGGAATCAGCGAGTGGATCC 420
Db 391 GTCAAGAAATTCAGCAGCTGAGAGAAACAGCTTGAATGTGCATATCAAGCGGAGACAG 450
Qy 421 AAAAGAACAGATGTTGCTTGAAGATCGCATCATCAGAGAGGGAAC 472
Db 451 AGAAGACGCACTGATGATGGAACAGGTGGAGGAACCTTCGCAGAAAGGAGC 502

RESULT 14

US-09-067-800-1
; Sequence 1, Application US/09067800
; Patent No. 6198024
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Ferrandiz, Cristina
; TITLE OF INVENTION: Seed Plants Characterized by Delayed
; TITLE OF INVENTION: Seed Dispersal
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/067,800
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1062 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:

; NAME/KEY: CDS
; LOCATION: 101..827
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1062
; OTHER INFORMATION: /note= "There is a poly(A) tail at
; OTHER INFORMATION: the end."
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..1062
; OTHER INFORMATION: /note= "Nucleotide and Deduced
; OTHER INFORMATION: Amino Acid Sequences of the AGL8 cDNA clone."
US-09-067-800-1

Query Match 15.6%; Score 141.6; DB 3; Length 1062;

Best Local Similarity 55.5%; Pred. No. 4.2e-32;
Matches 295; Conservative 0; Mismatches 234; Indels 3; Gaps 1;

Qy 1 ATGGGTCGTGGGAGATTGAGATTAAGAGGATTGAAATATCTAGAACCCAGAGGTCACT 60
Db 101 ATGGGAAGAGGTAGGTTTCAGCTGAAGAGGATAGAGAACAGATCAATAGGCAAGTTACT 160
Qy 61 TTCTGCAAGCGCGAAATGGTTTATTAAAGAGCGGTATGAATATATCAGTTCTTTGTGAT 120
Db 161 TTCTCAAGAGAGAGGTCTGGTTTCTCAAGAAAGCTCATGAGATCTCTGTCTCTGGAT 220
Qy 121 GCAGAAATGGCCCTCATCTCTTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACAC 180
Db 221 GCTGAGGTGCTCTCATCTCTCTCTCTCCAAAGGCAAACTCTTCGAATATTTCCACCGAC 280
Qy 181 ---ACCGTGAAGAGGAGGATTCAGAGGTACAGAGACTTCCGTTGACAAACACCCAGCA 237
Db 281 TCCTTGCATGGAGAGGATCTTGAACGCTATGTCGTATTATTCAGCAAAACAACCTT 340
Qy 238 GGGCGCATATCAGAGTCCAAATTTCTCAGTATTGGCAACAGGAGGCTGGTAAACTCAGACAA 297
Db 341 GTTGGCCGAGACGTTTCACAAAGTGAATTTGGTTCTTAGAACATGCTAAGCTCAAGGCA 400
Qy 298 CAGATTGACATTTGCAAAATGCAAAATAGGATTTGATGGTGACGGCTTACAGCTTTG 357
Db 401 AGAGTTGAGGTACTTTGAGAAGAACAAAGGAATTTTATGGGGGAAGATCTTTGATTCGTTG 460
Qy 358 AACATTAAAGGAACCTCAAGCAACTTGAGTTTCGACTTGAATAAGGAATCAGCCGAGTGCGA 417
Db 461 AGCTTGAAGAGGCTCCAAAGCTTGGAGCATCAGTCCATGATCAAGAGCATTTAGG 520
Qy 418 TCCAAAAAGAACGAGATGTTGCTTTGAAGAGATCGACATCATGCAAGAGAGGGAACACATA 477
Db 521 TCAAGAAAGAACCAAGCTATGTTCCGAATCCATATCTCGCTCCAGAAAGGATAAAGCC 580
Qy 478 CTTATCCAGGAGATGAGATTCTTTCGAGCAAGATAGCCAGTGTCAAGATA 529
Db 581 TTGCAAGATCAACAATTCGCTTCTCAAAAGATTAAAGGAGAGGAGAGAGA 632

RESULT 15

US-09-105-652-1
; Sequence 1, Application US/09105652
; Patent No. 6229068
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Martienssen, Robert
; APPLICANT: Ferrandiz, Cristina
; APPLICANT: Gu, Qing
; TITLE OF INVENTION: Method of Increasing Fruit Size in a Plant
; FILE REFERENCE: P-UD 3040
; CURRENT APPLICATION NUMBER: US/09/105,652
; CURRENT FILING DATE: 1998-06-26
; EARLIER APPLICATION NUMBER: 60/051,030
; EARLIER FILING DATE: 1997-06-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

```
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(826)
; FEATURE:
; NAME/KEY: polyA_signal
; LOCATION: (1061)
US-09-105-652-1

Query Match: 15.6%; Score 141.6; DB 3; Length 1062;
Best Local Similarity 55.5%; Pred. No. 4.2e-32;
Matches 295; Conservative 0; Mismatches 234; Indels 3; Gaps 1;

QY 1 ATCGGTCGTGGGAAGATTGAGATAAAGAGAGATTGAAAAATCTACGAACCGACGAGTCACT 60
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 TTCTGCAAGCGCCGAAATGTTTATTAAAGAGCGGTATGAATTTATCAGTCTCTTTGTGAT 120
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 161 TTCTCAAGAGAGAGCTGTGTTTCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 GCAGAAATGCGCCCTCATCGTCTTCTCCAGAGAGGAGACTTTATGAATTTGCCAACCCAC 180
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 221 GCTGAGGTTGCTCTCATCGTCTTCTCTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 ---AGCGTGAAGAGAGCGATTGAGAGGTACAGAGACTTGCCTTGACAACCAACCAACGGA 237
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 281 TCTTGCATGAGAGGATCTTGAACGCTATGATCGCTATTATATTCAGACAAACAACTT 340
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 238 GGGGCGGATATCAGAGTCCAAATCTCAGTATTGGCAACAGAGGCTCGTAAACTCAGACAA 297
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 341 GTTGCCCGACGCTTTCACAAAGTGAAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 400
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 298 CAGATTGACATTTTGCAAAATGCAATAGGCATTTGATGGTGACGGGCTTACAGCTTTG 357
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 401 AGAGTTGAGTACTTGAGAAAGCAAAAGGAATTTTATGGGGAAGATCTTGATTCGTTG 460
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 358 AACATTAAAGNACTCAAGCACTTGAGTTTCGACTTGAAAAAGGAATCAGCCGAGTSCGA 417
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 461 AGCTTGAAGAGCTCCAAAGCTTGGAGCATCAGTCGATCGATCAAGAGCATTAGG 520
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 418 TCCAAAAGAACCGAGATGTTGCTTGAAGAGATCGACATCATGAGAGAGAGGNAACACATA 477
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 521 TCAGAAAGAACCAAGCTATGTTTCGAATCCATATCTGCGCTCCAGAGAGAGGATAAAGCC 580
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 478 CTTATCCAGGAGATGAGATCTTTCGAGCAAGATAGCCGAGTGTGAGAATA 529
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 581 TTGCAAGATCAACAAATTCGCTTCTCATAAAGATTAAAGAGAGGAGGAGAGA 632
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: January 21, 2004, 22:19:47
Job time : 64.4877 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 15:50:10 ; Search time 3315.05 Seconds
(without alignments)
11217.584 Million cell updates/sec

Title: US-09-936-869-3
Perfect score: 909
Sequence: 1 atggctgtgtggaagattga.....cgcattgtattcttctgttg 909

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.in.v.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	909	100.0	922	8	AF023615 Pinus rad
2	885	97.4	1726	8	AF006210 Pinus res
3	782.8	86.1	1555	8	PMU69483 Picea maria
4	782.8	86.1	1642	8	PMU46582 Picea maria
5	778	85.6	1657	8	PMU69482 Picea maria
6	774.8	85.2	1753	8	PADAL2 X79280 P. abies dal
7	772.8	85.0	1212	8	PMU69484 U69484 Picea maria
8	455.6	50.1	1372	8	AY114304 A114304 Ginkgo bl
9	433	47.6	1164	8	AF492455 AF492455 Cycas ede
10	328.6	36.1	1154	8	AGN132209 AJ132209 Gnetum gn
11	303.8	33.4	912	8	AB050644 AB050644 Magnolia
12	279.6	30.8	942	8	AF373604 AF373604 Vitis vin
13	279	30.7	1009	8	AF234617 AF234617 Phalaenop
14	276.8	30.5	1142	8	AF234617 AF234617 Cucumis s
15	266	29.3	1001	8	AY178837 AY178837 Monardica
16	265.8	29.2	1160	8	AF265562 AF265562 Vitis vin
17	258.2	28.4	693	8	AF538966 AF538966 Gossypium
18	256.4	28.2	1048	8	AF035439 AF035439 Cucumis s
19	256.2	28.2	1131	8	SLSLM1 X80488 S. latifolia
20	255.6	28.1	959	6	AR138348 AR138348 Sequence
21	255.6	28.1	963	8	ATHAGLSA W55553 Arabidopsis
22	255	28.1	869	8	AB079260 AB079260 Agapanthu
23	254.6	28.0	1065	8	MDO251117 MDO251117 Malus dom
24	254.6	28.0	1125	6	AR003863 AR003863 Sequence
25	254.6	28.0	1125	8	TOBNAG1A L23925 Nicotiana t
26	254	27.9	1165	8	BPE252071 BPE252071 Betula p
27	253	27.8	1084	8	TOMTAG1A L36295 Lycopersico
28	251.4	27.7	1157	6	AX006357 AX006357 Sequence
29	251.4	27.7	1158	8	AF035438 AF035438 Cucumis s
30	251.4	27.7	1160	8	AF022379 AF022379 Cucumis s
31	251	27.6	1214	8	PHPMAD33 X72912 P. hybrida p
32	250.4	27.5	1061	8	AF194335 AF194335 Hyacinthu
33	250.2	27.5	1061	8	AF099937 AF099937 Hyacinthu
34	248.2	27.3	1297	6	AR227777 AR227777 Sequence
35	247.2	27.2	1179	8	AF027376 AF027376 Corylus a
36	245.4	27.0	1080	8	FCORFAG2 Z46612 P. ginseng g
37	245	27.0	895	6	ATHAGLSA M55550 Arabidopsis
38	245	27.0	895	6	AR138347 AR138347 Sequence
39	243.6	26.8	1037	8	AY086196 AY086196 Arabidops
40	243.6	26.8	1037	8	MDO251118 MDO251118 Malus dom
41	242.2	26.6	967	8	CSCUS1 X97801 C. sativus C
42	241.8	26.6	1065	8	AB025643 AB025643 Rosa rugo
43	241.6	26.6	916	8	AY098736 AY098736 Lycopersi
44	241.2	26.5	1000	8	AF022378 AF022378 Cucumis s
45	240.4	26.4	1073	8	S53900 S53900 ple=plena l

ALIGNMENTS

RESULT 1
AF023615

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AF023615
Pinus radiata MADS box protein mRNA, complete cds.
922 bp mRNA linear
PLN 26-JAN-1999

Pinus radiata (Monterey pine)
Pinus radiata

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 922)
Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
Jun-Jun Liu and G. K. Podila.
Direct Submission
Submitted (09-SEP-1997) Department of Biological Sciences, Michigan


```
Technological University, Houghton, MI 49931, USA
Location/Qualifiers
1..922
/organism="Pinus radiata"
/mol_type="mRNA"
/db_xref="taxon:3347"
14..682
/codon_start=1
/product="MADS box protein"
/protein_id="AAD09342.1"
/db_xref="GI:4103486"
/translation="MGRGKIEIKRIENTNRQVTFCKRRNGLLKAVELSVLCDAEVA
LIVFSRGLYEFANHSKRTIERYKTCVNNHGGAI:SESNQYQWQEGAKRQQID
ILQANRHLMGDLTALNLIKQLEVRLEKIGSRVSKNEMLEEDIMORREHIL
IQENELRSKIAECQNSHNTNMLSAPEYDALPAPDSRNLFLANLIDAAHYAHQEQTT
LQLG"
BASE COUNT 303 a 174 c 218 g 227 t
ORIGIN
Query Match 100.0%; Score 909; DB 8; Length 922;
Best Local Similarity 100.0%; Pred. No. 4.3e-238;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGGTCGCGGAGATTGAGTAAAGAGGATTGAAATCTACGAAACCGCAGGTCACT 60
DB 14 ATGGGTCGCGGAGATTGAGTAAAGAGGATTGAAATCTACGAAACCGCAGGTCACT 73
QY 61 TTCTGCAAGCGCGGAATGGTTTATTAAGAAGGGCGTATGAATATCACTTTCTTGAT 120
DB 74 TTCTGCAAGCGCGGAATGGTTTATTAAGAAGGGCGTATGAATATCACTTTCTTGAT 133
QY 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGGAGGAGACTTATGAATTCGCCAACAC 180
DB 134 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGGAGGAGACTTATGAATTCGCCAACAC 193
QY 181 AGCGTGAAGGAGGAGGAGTACAGAGGTACAAAGAGCTTGGGTTGACAAACCAACCGAGGG 240
DB 194 AGCGTGAAGGAGGAGGAGTACAGAGGTACAAAGAGCTTGGGTTGACAAACCAACCGAGGG 253
QY 241 GCAGATCAGAGTCCCAATCTCAGTATTGGCAACAGAGAGCTGGTAACTCAGAACAG 300
DB 254 GCAGATCAGAGTCCCAATCTCAGTATTGGCAACAGAGAGCTGGTAACTCAGAACAG 313
QY 301 ATTGACATTTTGCAAAATGCAATAGGATTTGATGGGTGACCGGCTTACAGCTTTGAC 360
DB 314 ATTGACATTTTGCAAAATGCAATAGGATTTGATGGGTGACCGGCTTACAGCTTTGAC 373
QY 361 ATTAAGGAACCTCAAGCAACTTGAGGTTGCACTTGAAGAAAGGATCAGCGAGTGGCATCC 420
DB 374 ATTAAGGAACCTCAAGCAACTTGAGGTTGCACTTGAAGAAAGGATCAGCGAGTGGCATCC 433
QY 421 AAAAGNACGAGATGTTGTTGAAGAGATCGACATCATGAGAGAGGGAACACATACCT 480
DB 434 AAAAGNACGAGATGTTGTTGAAGAGATCGACATCATGAGAGAGGGAACACATACCT 493
QY 481 ATCCAGAGAGATGAGATTCTTCGAGCAAGATAGCCGAGTGTGAGATAGCCACACAG 540
DB 494 ATCCAGAGAGATGAGATTCTTCGAGCAAGATAGCCGAGTGTGAGATAGCCACACAG 553
QY 541 AACATGTTATCAGTCCGGAATATGATGCACTGCCCGCATTCGACTCTCGAATTTCTTA 600
DB 554 AACATGTTATCAGTCCGGAATATGATGCACTGCCCGCATTCGACTCTCGAATTTCTTA 613
QY 601 CATGCAATCTAATCGATGCGGCCCATCACTATGACATCAGGAACAAACACCGTTTCA 660
DB 614 CATGCAATCTAATCGATGCGGCCCATCACTATGACATCAGGAACAAACACCGTTTCA 673
QY 661 CTTGGCTGAAGCGGTGAGCGCTTAAATCTCAATCAAGGACCCGGAATATG 720
DB 674 CTTGGCTGAAGCGGTGAGCGCTTAAATCTCAATCAAGGACCCGGAATATG 733
QY 721 CTAGTAACTTGAATGAGATTCAGTTCGAAATATTCGAGGCAAGAGGCTGGTAACTCAGACAG 780
DB 721 CTAGTAACTTGAATGAGATTCAGTTCGAAATATTCGAGGCAAGAGGCTGGTAACTCAGACAG 790
```

```
734 CTAGTAACTTGAATGAGATTCAGTTCGAAATATTCGAGGCAAGAGGCTGGTAACTCAGACAG 793
781 GATAGCTCCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
794 GATAGCTCCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 853
841 ATAGCTGTTTGAACAAAGAAATCAACATATGATGATGATGATGATGATGATGATGATGATGAT 900
854 ATAGCTGTTTGAACAAAGAAATCAACATATGATGATGATGATGATGATGATGATGATGATGAT 913
901 TTCTCTGTTG 909
914 TTCTCTGTTG 922
RESULT 2
AF006210 1726 bp mRNA linear PLN 05-JAN-1999
LOCUS Pinus resinosa MADS box transcription factor mRNA, complete cds.
DEFINITION AF006210
ACCESSION AF006210 GI:4101709
VERSION AF006210.1
KEYWORDS Pinus resinosa
SOURCE Pinus resinosa
ORGANISM Pinus resinosa
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 1726)
Liu, J.-J. and Podlita, G.K.
Direct Submission
Submitted (01-JUN-1997) Department of Biological Sciences, Michigan
Technological University, Houghton, MI 49931, USA
Location/Qualifiers
1..1726
/organism="Pinus resinosa"
/mol_type="mRNA"
/db_xref="taxon:54921"
/sex="female"
/dev_stage="immature cones"
/notes="Lambda Triplex cDNA library"
475..1143
/codon_start=1
/product="MADS box transcription factor"
/protein_id="AAD01266.1"
/db_xref="GI:4101710"
/translation="MGRGKIEIKRIENTNRQVTFCKRRNGLLKAVELSVLCDAEVA
LIVFSRGLYEFANHSKRTIERYKTCVNNHGGAI:SESNQYQWQEGAKRQQIE
ILQANRHLMGDLTALNLIKQLEVRLEKIGSRVSKNEMLEEDIMORREHIL
IQENELRSKIAECQNSHNTNMLSAPEYDALPAPDSRNLFLANLIDAAHYAHQEQTT
LQLG"
BASE COUNT 527 a 346 c 382 g 471 t
ORIGIN
Query Match 97.4%; Score 885; DB 8; Length 1726;
Best Local Similarity 98.3%; Pred. No. 1.7e-231;
Matches 894; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 ATGGTCTGCGGAGATTGAGTAAAGAGGTTGAAATCTACGAAACCGCAGGTCACT 60
DB 475 ATGGTCTGCGGAGATTGAGTAAAGAGGTTGAAATCTACGAAACCGCAGGTCACT 534
QY 61 TTCTGCAAGCGCGGAATGGTTTATTAAGAAGGGCGTATGAATATCACTTTCTTGAT 120
DB 535 TTCTGCAAGCGCGGAATGGTTTATTAAGAAGGGCGTATGAATATCACTTTCTTGAT 594
QY 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGGAGGAGACTTATGAATTCGCCAACAC 180
DB 595 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGGAGGAGACTTATGAATTCGCCAACAC 654
QY 181 AGCGTGAAGGAGGAGGATGAGAGGTACAAAGAGCTTGGGTTGACAAACCAACCGAGGG 240
DB 655 AGCGTGAAGGAGGAGGATGAGAGGTACAAAGAGCTTGGGTTGACAAACCAACCGAGGG 714
QY 241 GCGATATCAGAGTCCCAATCTCAGTATTGGCAACAGGAGGCTGGTAACTCAGACACAG 300
```

Db	715	CGGATTCAGATCCCAATCTCAGTATTTGGCAACAGGAGCTGGTAACTCAGACAACAG	774
Qy	301	ATTGACATTTTGCAGAAATCAATAGGCAATTTGATGGGTGAGCGCTTACAGCTTTGAAC	360
Db	775	ATTGAAATTTTGCAGAAATCAATAGGCAATTTGATGGGTGAGCGCTTACAGCTTTGAAC	834
Qy	361	ATTAGGAATCAAGCAATCTGAGGTTTGCACCTTGAAGAAAGGAATCAGCGGAGTGGATCC	420
Db	835	ATCAGGAATCAAGCAATCTGAGGTTTGCACCTTGAAGAAAGGAATCAGCGGAGTGGATCC	894
Qy	421	AAAAGAAACAGATGTTGTTGAAGAGATCGACATCATGCGAGAGAGGAACACATCTT	480
Db	895	AAAAGAAACAGATGTTGTTGAAGAGATCGACATCATGCGAGAGAGGAACACATCTT	954
Qy	481	ATCCAGGAGATGAGATTTCTTGCAGCAAGATAGCGGAGTGCAGATAGCCACAACACG	540
Db	955	ATCCAGGAGATGAGATTTCTTGCAGCAAGATAGCGGAGTGCAGATAGCCACAACACG	1014
Qy	541	AACATGTTATGAGTCCGGAATATGATGCACTGCCCGCATTCGATCTCGAAATTTCTA	600
Db	1015	AACATGTTATGAGTCCGGAATATGATGCACTGCCCGCATTCGATCTCGAAATTTCTA	1074
Qy	601	CATGCAAACTCAATGATCGGCGCCCATCTACTATGCATCGAGAAACACACGCTTCAG	660
Db	1075	CATGCAAACTCAATGATCGGCGCCCATCTACTATGCATCGAGAAACACACGCTTCAG	1134
Qy	661	CTTGCTGAAAGCTTGAAGCGGTGGAGCTTAAATCTCAATCAAGGCAACCCGAAATATG	720
Db	1135	CTTGCTGAAAGCTTGAAGCGGTGGAGCTTAAATCTCAATCAAGGCAACCCGAAATATG	1194
Qy	721	CTAGTAACCTTCAATGAGATTCAGATCGAATATTTGGAGGCAAGACACATGGAAGA	780
Db	1195	CTAGTAACCTTCAATGAGATTCAGATCGAATATTTGGAGGCAAGACACATGGAAGA	1254
Qy	781	GATAGCTCTAGTATGAATATGATTTATGATTAATTAACATATGTTTGTGAGCTTTAAT	840
Db	1255	GATAGCTCTAGTATGAATATGATTTATGATTAATTAACATATGTTTGTGAGCTTTAAT	1314
Qy	841	ATAGCTGTTTGAACAAAGATTAACATATTAACATATTAACATATGTTTGTGAGCTTTAAT	900
Db	1315	ATAGCTGTTTGAACAAAGATTAACATATTAACATATTAACATATGTTTGTGAGCTTTAAT	1374
Qy	901	TTTCTGTTG 909	
Db	1375	TTTCTGTTG 1383	

RESULT 3
 PMU69483
 LOCUS
 DEFINITION
 Picea mariana AGAMOUS-like MADS-box transcriptional factor SMADS42C
 mRNA, complete cds.
 ACCESSION
 VERSION
 U69483.1 GI:4033722
 KEYWORDS
 SOURCE
 Picea mariana
 ORGANISM
 Picea mariana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
 REFERENCE
 AUTHORS
 Rutledge, R., Regan, S., Nicolas, O., Fobert, P., Cote, C., Bosnich, W.,
 Kauffeldt, C., Sunchara, G., Seguin, A. and Stewart, D.
 TITLE
 Characterization of an AGAMOUS homologue from the conifer black
 spruce (*Picea mariana*) that produces floral homeotic conversions
 when expressed in *Arabidopsis*
 JOURNAL
 MEDLINE
 98451867
 PUBMED
 9778845
 REFERENCE
 2 (bases 1 to 1555)
 AUTHORS
 Rutledge, R.G., Kauffeldt, C., Fobert, P., Cote, C., Bosnich, W. and
 Stewart, D.
 TITLE
 Direct Submission

JOURNAL	Submitted (04-SEP-1996) Canadian Forest Service, Laurentian Forestry Centre, 1055 du P.E.P.S., Staince-Foy, PQ G1V 4C7, Canada		
FEATURES	Location/Qualifiers		
source	1..1555		
	/organism="Picea mariana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3335"		
	/sex="female"		
	/tissue_type="cones"		
	/dev_stage="just prior to pollination"		
	329..997		
CDS	/codon_start=1		
	/product="AGAMOUS-like MADS-box transcriptional factor SMAD542C"		
	/protein_id="AAC97158.1"		
	/db_xref="GI:4033723"		
	/translation="MGRKIEIKRIENTNTROVTFCKRRNLKKAYELSVLCDAEVA LIVPSRGLYEFANSHVKRTIERVKTCDVNNHGAISNSQYQWQAGKROQIE ILQNRHLMGDLTALNKEIKQLSEVLEKGIQVRSKNEMLEIDIMORREHIL IQENELRSKIAECQNSHTSMLSAFEDALPAFDSRNFHLANLIDAAHYHQEQT LQIG"		
BASE COUNT	476 a	297 c	348 g 434 t
ORIGIN			
Query Match	86.1%; Score 782.8; DB 8; Length 1555;		
Best Local Similarity	93.1%; Pred. No. 1.8e-203;		
Matches	855; Conservative 0; Mismatches 52; Indels 11; Gaps 3;		
QY	1	ATGGGTCGTGGGAGATTGAGATAAGAGGATTGAAATACTACGAAACCGACAGGTCAC	60
Db	329	ATGGGCGGTGGGAGATTGAGATAAGAGGATTGAAATACTACGAAACCGACAGGTCAC	388
QY	61	TTCTGCAAGCGCCGAATGCTTTATTAAAGAGGCGGTATGAATATTCAGTTCTTTGTGAT	120
Db	389	TTCTGTAAGCGCGGAATGCTTTATTGAAGAGGCGGTATGAATATTCAGTTCTTTGTGAT	448
QY	121	GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGGACTTTATGAATTTGCCAACCCAC	180
Db	449	GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGGACTTTATGAGTTTGCACACCCAC	508
QY	181	AGCGTGAAGAGGACGATTGAGAGGTACAAGAACACTTTCGTTTGCACAAACCAACCGAGGG	240
Db	509	AGCGTGAAGAGAACGATTGAGAGGTACAAGAACACTTTCGTTTGCACAAACCAACCGAGGG	568
QY	241	GCGATATCAGATGCCAATTTCTCAGTATTGGCAACAGAGGCTGGTAACTCAGACAACACAG	300
Db	569	GCGATTTCAAGTCCCAATTTCTCAGTATTGGCAACAGAGGCTGGTAACTCAGACAACACAG	628
QY	301	ATTGACATTTTGCAGAAATGCAATAGGCAATTTGATGGGTGACGGGCTTACAGCTTTGAAC	360
Db	629	ATTGAAATTTTGCAGAAATGCAATAGGCAATTTGATGGGTGACGGGCTTACAGCTTTGAAC	688
QY	361	ATTAAGGAATCAGCAACTTGGAGTTTCGACTTTGAAGAAAGGAATCAGCCGAGTGGCATCC	420
Db	689	ATCAAGGAATCAGCAACTTGGAGTTTCGACTTTGAAGAAAGGAATCAGCCGAGTGGCATCC	748
QY	421	AAAAAGAACGAGATGTTGTTGAAGAGATCGACATCATGCGAGAGGGAACACATCTT	480
Db	749	AAGAAGAACGAGATGTTGTTGAAGAGATCGACATCATGCGAGAGGGAACACATCTT	808
QY	481	ATCCAGAGGAATGAGATTTCTTCCAGCAAGATAGCCGAGTGCAGATAGCCACAACACG	540
Db	809	ATCCAGAGGAATGAGATTTCTTCCAGCAAGATAGCCGAGTGCAGATAGCCACAACACG	868
QY	541	AACATGTTATCAGCTCCGGAATATGATGCACTGCCCGCATTCGACTTCGAAATTTCTTA	600
Db	869	AGCATGCTATCAGCTCCGGAATATGATGCACTGCCCGCATTCGACTTCGAAATTTCTTA	928
QY	601	CATGCAAACTCAATCGATCGGCGCCCATCTACTATGCATCAGGAAACAAACACGCTTCAG	660
Db	929	CATGCAAACTCAATCGATCGGCGCCCATCTACTATGCATCAGGAGCAAAACACGCTTCAG	988
QY	661	CTTGGCTGAACGTTGAAGCGGTGGAGCTTAAACTCAATCAAGGCAACCCGAAATAATG	720


```

KEYWORDS      Picea mariana
SOURCE        Picea mariana
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
REFERENCE     Rutledge,R., Regan,S., Nicolas,O., Fobert,P., Cote,C., Bosnich,W.,
AUTHORS       Kauffeldt,C., Sunohara,G., Seguin,A. and Stewart,D.
TITLE         Characterization of an AGAMOUS homologue from the conifer black
              spruce (Picea mariana) that produces floral homeotic conversions
              when expressed in Arabidopsis
JOURNAL       Plant J. 15 (5), 625-634 (1998)
MEDLINE       98451867
PUBMED        9778845
REFERENCE     Rutledge,R.G., Kauffeldt,C., Fobert,P., Cote,C., Bosnich,W. and
AUTHORS       Stewart,D.
TITLE         Direct Submission
JOURNAL       Submitted (04-SEP-1996) Canadian Forest Service, Laurentian
              Forestry Centre, 1055 du P.E.P.S., Sainte-Foy, PQ G1V 4C7, Canada
FEATURES      Location/Qualifiers
              1..1657
               /organism="Picea mariana"
               /mol_type="mRNA"
               /db_xref="taxon:3335"
               /sex="female"
               /tissue_type="cones"
               /dev_stage="just prior to pollination"
               457..1125
               /codon_start=1
               /product="AGAMOUS-like MADS-box transcriptional factor
              SAG1a"
               /protein_id="AAC97157.1"
               /db_xref="GI:4033721"
               /translation="MGRKLEIKRIENTNNQVTFCKRBNGLLKAYELSVLCDAEVA
              LIVFSRGRLEYPANSHVKTIRTYKKTCDNNHGAISENSQVQWQKRLQRIE
              ILQNAHRHLMGDLNLTAKELKQLEKIGRVRKXNEMLEEDIMORREHIL
              IQENEILRSKIARCONSHNTNLSAPEYDALPFDNRNLFHANLIDAAHHYHQQT
              LQUG"
CDS           497 a 336 c 362 g 462 t
              BASE COUNT 497 a 336 c 362 g 462 t
              ORIGIN
              Query Match 85.6%; Score 778; DB 8; Length 1657;
              Best Local Similarity 92.8%; Pred. No. 3.7e-202;
              Matches 852; Conservative 0; Mismatches 55; Indels 11; Gaps 3;
QY 1 ATGGTCTGTCGGAAGATTGAGATAAGAGGATTGAAATGAAAATGCTACGAAACGACAGGTCAC 60
DB 457 ATGGCCCGTGGAGATTGAGATAGAGGATTGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 516
QY 61 TTCTGCAAGCCCGAAATGTTTATTAAAGAGGCGGTATGAAATATCAGTTCTTTGTGAT 120
DB 517 TTCTGTAAGCCCGAAATGTTTATTGAGAGGCGGTATGAAATATCAGTTCTTTGTGAT 576
QY 121 GCAGAGTGCCTCATCGTCTCTCCAGCAGAGGAGACTTTATGAAATTTGCCAACCCAC 180
DB 577 GCAGAGTGCCTCATCGTCTCTCCAGCAGAGGAGACTTTATGAGTTTGCCAACCCAC 636
QY 181 AGCGTGAAGAGGAGGATTGAGAGGTACAAGAGACTTCGTTGACAAACACCGAGGG 240
DB 637 AGCGTGAAGAGAGGATTGAGAGGTACAAGAGACTTCGTTGACAAACACCGAGGG 696
QY 241 GCGATATCAGATCCAAATTCAGTATTTGCAACAGAGGCTGGTAACTCAGACACAG 300
DB 697 GCGATTTCAGATCCAAATTCAGTATTTGCAACAGAGGCTGGTAACTCAGACACAG 756
QY 301 ATTACATTTTGCATAATGCAATAGGCAATTTGATGGGTGACGGCTTACAGCTTTGAAC 360
DB 757 ATTGAATTTTACAATATGCAATAGGCAATTTGATGGGTGATGGCTTACAGCTTTAAAC 816
QY 361 ATTAAGGAACCTCAAGCACTTGAGGTTGCGATCTGAAAAGAAATCAGCCGAGTGCATCC 420
DB 817 ATCAAGGAACCTCAAGCACTTGAGGTTGCGATCTGAAAAGAAATCAGCCGAGTGCATCC 876

```

```

QY 421 AAAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATGCGAGAGAGGGACACATCTT 480
DB 877 AAGAGAACGAGATGTTGCTTGAAGAGATCGACATCATGCGAGAGAGGGACACATCTT 936
QY 481 ATCCAGAGAGATGAGATTCTTCGAGCAAGATAGCCGAGTGTGAGATAGCCACACACG 540
DB 937 ATCCAGAGAGATGAGATTCTTCGAGCAAGATAGCCGAGTGTGAGATAGCCACACACG 996
QY 541 AACATGTTATCAGCTCCGGAATATGATGCACTGCCGCCATTCGACTCTCGAAATTTCTTA 600
DB 997 AACATGCTATCAGCTCCGGAATATGATGCACTGCCGCCATTCGACTCTCGAAATTTCTTA 1056
QY 601 CATGCAAAATCTAATCGATGGGCCCATCATCATGACATCAGGAAACAAACACGCTTCAG 660
DB 1057 CATGCAAAATCTAATCGATGGGCCCATCATCATGACATCAGGAAACAAACGCTTCAG 1116
QY 661 CTTGCGTGAAGTTGAAGCGGTGGACGCTTAAACTCAATCAAGGCAACCGGAAATATG 720
DB 1117 CTTGCGTGAAGTTGAAGCGGTGGATGCTTAAACTCAATCAAGGCTCCTTTGAAATATG 1176
QY 721 CTAGTAACTTGAATGAGAT-----TCAGAGTGGAAATATTTGCGAGGCAAGAGCAC 771
DB 1177 -TTGTAGCTTTGAATGAGATTCCTCATGATCCCATGATCCCATGATTTGTGAGGCAAAATCAC 1235
QY 772 AATGGAAGAGATAGCTCTCTAGTATGAATATGATGATTTATGATATTAACATATGTTGTCA 831
DB 1236 AATGAAGGAGATAGCTCTCTAGTATAAATATGATGATTTATGATATTAACATATGTTGTCA 1295
QY 832 GCTTTAATATAGCTGTTTGAACAAAGATACACATATAGCTAGTATTTTGTGCG 891
DB 1296 GTTTTATATATAGCTGTTTGAACAAAGATAGAACATATATGCTA-TTGTTCCTCGCG 1354
QY 892 CATGTTATCTTCTGTTG 909
DB 1355 CATTTAGCTTTGCTG 1372
RESULT 6
PADAL2      PADAL2      1753 bp      mRNA      linear      PLN 28-FEB-1995
LOCUS       P.abies dal2 mRNA.
DEFINITION  X79280
ACCESSION   X79280
VERSION     X79280.1 GI:695687
KEYWORDS    DAL2 gene.
SOURCE      Picea abies (Norway spruce)
ORGANISM   Picea abies
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Picea.
REFERENCE   1
AUTHORS     Tandré,K., Albert,V.A., Sundas,A. and Engström,P.
TITLE       Conifer homologues to genes that control floral development in
            angiosperms
JOURNAL     Plant Mol. Biol. 27 (1), 69-78 (1995)
MEDLINE     95170009
PUBMED      7865797
REFERENCE   2 (bases 1 to 1753)
AUTHORS     Tandré,K.
TITLE       Direct Submission
JOURNAL     Submitted (16-MAY-1994) K. Tandré, Uppsala University, Dept of
            Physiological Botany, Villavägen 6, 752 36 Uppsala, SWEDEN
FEATURES    Location/Qualifiers
            1..1753
             /organism="Picea abies"
             /mol_type="mRNA"
             /db_xref="taxon:3329"
             /clone="3ae"
             /tissue_type="female cone"
             /clone_lib="cDNA library spruce female; SF"
             /dev_stage="before formation of ovule"
            1..1753
             /gene="dal2"
             424..1092
            gene
            CDS

```

		/gene="dal2"			
		/codon_start=1			
		/product="DAL2 protein"			
		/protein_id="CAA5867.1"			
		/db_xref="GI:695688"			
		/db_xref="SPTREMBL:Q40766"			
		/translation="MWRGKLEIKRITENTNROVTFCKRNLGLKKAYELSVLCDAEVALIVFSRGLYFAHFASVKRTIERKYKTCVDNNHGVISESQVWQEGAKLRQGLEILONAHMIDGLTALNTKEIKGLVEVRLEKGIQRVRSKNMLLEEDIMQRRRHILIOENELIRSKIAEQCONSHNTNMLSAPEYDALPAFDSRNPFLHANLIDAAHHYAHQSQTTILQLG"			
misc_feature		427..594			
		/gene="dal2"			
		/note="MADS box"			
misc_feature		697..891			
		/gene="dal2"			
		/note="K box"			
BASE COUNT	529 a	340 c	388 g	496 t	
ORIGIN					
	Query Match	85.2%	Score 774.8;	DB 8;	Length 1753;
	Best Local Similarity	92.8%;	Pred. No. 2.8e-201;		
	Matches 850;	Conservative 0;	Mismatches 57;	Indels 11;	Gaps 3;
QY	1	ATGGGTCGTGGGAAGATTGAGATAAAGAGAGATTGAAAAATACTACGAACCGACAGGTCAC	60		
Db	424	ATGGCCGCTGGGAAGATTGAGATAAAGAGATTGAGATACTACGAACCGACAGGTCAC	483		
QY	61	TTCTGCAAGCGCCGAATGTTTTATTAAGAAGCGGTATGAAATATCAGTTCTTTGTGAT	120		
Db	484	TTCTGTAAGCGCCGAATGTTTTATTAAGAAGCGGTATGAAATATCAGTTCTTTGTGAT	543		
QY	121	GCAGAAGTGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTATGAAATTTGCCAACCC	180		
Db	544	GCAGAAGTGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTATGAGTTTGCCAACCC	603		
QY	181	AGCGTGAAGAGGACGATTGAGAGGTAAGAAGACTTTCGTTGACCAACACCGCGAGGG	240		
Db	604	AGCGTGAAGAGAACCGATTGAGAGGTAAGAAGACTTTCGTTGACCAACACCGCGAGGG	663		
QY	241	GCATATTCAGAGTCCCAATTCCTCAGTATGGCAACAGGAGGCTGGTAAACTCAGACAA	300		
Db	664	GTGATTTCCGAGTCCCAATTCCTCAGTATGGCAACAGGAGGCTGGTAAACTCAGACAA	723		
QY	301	ATTGACATTTTGAAAAATGCAAAATAGGCATTTGATGGGTGACGGGCTTACGTTTGAAC	360		
Db	724	ATTGAAATTTTGAAAAATGCAAAATAGGCACCTTGAATGGGTGACGGGCTTACGTTTGAAC	783		
QY	361	ATTGAAGCACTCAGAGCAACTTGAGGTTGCACTTGCAAAAAGAAATCAGCCGAGTCGGATCC	420		
Db	784	ATCAAGGACCTCAGAGCACTTGAGGTTGCACTTGCAAAAAGAAATCAGCCGAGTCGGATCC	843		
QY	421	AAAAAAGAACGAGATGTTGCTTTGAAGAGATCGACATCATGCAAGAAAGGGAACACATCTT	480		
Db	844	AAGAAGAACGAGATGTTGCTTTGAGAGATCGACATCATGCAAGAAAGGGAACACATCTT	903		
QY	481	ATCAGGAGATGAGATTTCTTCGCAGCAAGATAGCCGAGTGTGAGAAATAGCCACAAACG	540		
Db	904	ATCCAGGAGATGAGATTTCTTCGCAGCAAGATAGCCGAGTGTGAGAAATAGCCACAAACG	963		
QY	541	AACATGTTATCAGCTCCCGGAATATGATGCACTGCCCGCAATTCGATCTCGAAATTTCTTA	600		
Db	964	AACATGCTATCAGCTCCCGGAATATGATGCACTGCCCGCAATTCGATCTCGAAATTTCTTA	1023		
QY	601	CATGCAAAATCTAATCGATCGGCCCATCATTATGCACATCAGGACCAACACGCTTCAG	660		
Db	1024	CATGCAAAATCTAATCGATCGGCCCATCATTATGCACATCAGGACCAACACGCTTCAG	1083		
QY	661	CTTGGCTCAACGCTTGAAGCGGTGGACGCTTAAAACTCAATCAAGGCAACCGAAAAATATG	720		
Db	1084	CTTGGCTCAAGGTTGAAGTCGTGGATGCTTAAAACTCAATCAAGGTCCTTGGAAAAATATG	1143		
QY	721	CTAGTAAACCTTGAATGAGAT-----TCAGATCGAAATATTTCCGAGGCGCAGAGCAC	771		

QY 71 GCCGAATCGTTTATTAAAGAGCGGTATGAATTATCAGTTCTTTTGTGATCGAGAGTGG 130
Db 61 GCCGAATCGTTTATTAAAGAGCGGTATGAATTATCAGTTCTTTTGTGATCGAGAGTGG 120
QY 131 CCTCATCGTCTTCTCCAGCAGGAGAGACTTTTAAATTTCCCAACACACAGCGTGAAGA 190
Db 121 CCTCATCGTCTTCTCCAGCAGGAGAGACTTTTAAATTTCCCAACACACAGCGTGAAGA 180
QY 191 GCAGATTGAGAGGTACAGAGACTTGGCTTTGACAAACACACAGGAGGGGGATATCAG 250
Db 181 GAACGATTGAGAGGTACAGAGACTTGGCTTTGACAAACACACAGGAGGGGGATATCAG 240
QY 251 AGTCCAAATCTCAGTATTCGCAACAGAGCGCTGGTAAACTCAGACACAGATTGACATT 310
Db 241 AGTCCAAATCTCAGTATTCGCAACAGGAGCGCTGGTAAACTCAGACACAGATTGAAATTT 300
QY 311 TGCAAAATGCAATAGGCAATTTGATGGGTACGGCTTTACAGCTTTGAACTTTAAAGAAC 370
Db 301 TGCAAAATGCAATAGGCAATTTGATGGGTACGGCTTTACAGCTTTTAAACATCAAGGAAC 360
QY 371 TCAAGCAACTTGAGGTTCGACTTTGAAAAAGGAATCAGCGAGTGGGATCCAAAAAGAACG 430
Db 361 TCAAGCAACTTGAGGTTCGACTTTGAAAAAGGAATCAGCGAGTGGGATCCAAAAAGAACG 420
QY 431 AGATGTTGCTTGAAGAGATCGACATCATCGAGAGAGGGAACACATATCTATCCAGGAGA 490
Db 421 AGATGTTGCTTGAAGAGATCGACATCATCGAGAGAGGGAACACATATCTATCCAGGAGA 480
QY 491 ATGAGATTCTTCGACCAAGATAGCGGAGTTCAGATAGGACAGACACAGACATGTTAT 550
Db 481 ATGAGATTCTTCGACCAAGATAGCGGAGTTCAGATAGGACAGACACAGACATGCTAT 540
QY 551 CAGCTCCGAATATGATGCACTGCCGCGATTCGACTCTCGAAATTTCTTACATGCAAAATC 610
Db 541 CAGCTCCGAATATGATGCACTGCCGCGATTCGACTCTCGAAATTTCTTACATGCAAAATC 600
QY 611 TAATCGATGGGCGCCATCTATGACATCAGAGCAAAACACAGCTTCAGCTGGCTGAA 670
Db 601 TAATCGATGGGCGCCATCTATGACATCAGAGCAAAACACAGCTTCAGCTGGCTGAA 660
QY 671 CGTTGAAGCGGTGGAGCTTAAACTCAATCAAGGCAACCGCAAAATATGCTAGTAACCT 730
Db 661 GGTGAAGTGGTGGATGCTTAAACTCAATCAAGGCTCTTGAATATG-TTGTAGCT 719
QY 731 TGAATCAGAT-----TCAGATCGAAATATTGCGAGGCAAGCAACATGGAAGAG 781
Db 720 TGAATCAGATTCCTATGGATCCAGTCCGAATATTGCGAGGCAAAATCAATGGAAGAG 779
QY 782 ATAGCTCCTAGTATGATGATGATTTATGATATTAACATATGTTTGTGAGCTTTAAATA 841
Db 780 ATAGCTCCTAGTATGATGATGATTTATGATATTAACATATGTTTGTGAGCTTTATATA 839
QY 842 TAGCTGTTTGAACAAAGAAATACACATATAGCTAGTATTTTTTTGGCGCATGTTATCT 901
Db 840 TAGCTGTTTGAACAAAGAAATAGACATATAGCTA-TTGTGTTTCTCGGCAATTTAGCT 898
QY 902 TTCTGTTG 909
Db 899 TTCTGTTG 906

RESULT 8
AY114304
LOCUS AY114304 1372 bp mRNA linear PLN 29-APR-2003
DEFINITION Ginkgo biloba AGAMOUS-like MADS-box transcription factor (GBM5)
mRNA, complete cds.
ACCESSION AY114304
VERSION AY114304.1 GI:30230269
KEYWORDS Ginkgo biloba (maidenhair tree)
SOURCE Ginkgo biloba
ORGANISM Ginkgo biloba
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Ginkgoophyta; Ginkgoales; Ginkgoaceae; Ginkgo.

REFERENCE 1 (bases 1 to 1372)
AUTHORS Jager,M., Hassanin,A., Manuel,M., Le Guyader,H. and Deutsch,J.
TITLE MADS-Box Genes in Ginkgo biloba and the Evolution of the AGAMOUS Family
JOURNAL Mol. Biol. Evol. 20 (5), 842-854 (2003)
PUBMED 12679535
REFERENCE 2 (bases 1 to 1372)
AUTHORS Jager,M., Hassanin,A., Manuel,M., Le Guyader,H. and Deutsch,J.
TITLE Direct Submission
JOURNAL Submitted (17-May-2002) Universite Paris VI, 9 quai Saint Bernard, Paris 75252 cedex05, France
FEATURES
Location/Qualifiers
source
1. .1372
/organism="Ginkgo biloba"
/mol_type="mRNA"
/db_xref="taxon:3311"
/feature_type="stamen"
1. .1372
/gene="GBM5"
1. .324
/gene="GBM5"
325..990
/gene="GBM5"
/codon_start=1
/product="AGAMOUS-like MADS-box transcription factor"
/protein_id="AA076208.1"
/db_xref="GI:30230270"
/translation="MGRKIEIKRIENTTNQVTFCKRRNGLLKKAYELSVLCDAEVA
LIVFSRGRILYEPANNSKVTIDRYKTKADNSGGGAISECNQYWOEAKRQOQID
ILQANPHLMGDALTSLSUKELQLEIRLGRGISRVSCKNMLLEEIEIMORREHIL
LASENFURTKIAECSSQNANMLPGPEFDALPGFDSRHLFASIMDAHYHAQQDQAL
QLG"

3'UTR 991..1372
BASE COUNT 439 a 270 c 275 g 388 t
ORIGIN

Query Match 50.1%; Score 455.6; DB 8; Length 1372;
Best Local Similarity 73.1%; Pred. No. 9.4e-114;
Matches 677; Conservative 0; Mismatches 219; Indels 30; Gaps 6;

QY 1 ATGGGTCGTGGAGATTGAGATAAGAGGATTGAAATATCTACGAACCGACAGGTCACCT 60
Db 325 ATGGGTCGTGGAGATTGAGATAAGAGGATTGAGAACTACGAACCGACAGGTCACCT 384
QY 61 TTCTGCAAGCGCCGAAATGGTTTATTAAAGAGCGGTATGAATATCAGTTCCTTTGAT 120
Db 385 TTCTGCAAGCGTCGAAATGGTCTGTTGAGAGAAAGCGGTATGAATATCCTGCTTTGAT 444
QY 121 GCAGAAGTGGCCCTCATCTCTCTCCAGCAGAGGAGACTTTTATGAATTTGCCAACAC 180
Db 445 GCGAAGTGGCT 504
QY 181 AGCGTGAAGAGGAGGATGAGAGGTACAGAAGACTTGGCTTGAACAACACCGAGGG 240
Db 505 AGTGTGAAGAGAAACAATCGACAGATACAGAAGACTTGGCTGACAACTCTCGAGGTGA 564
QY 241 GCGATATCAGAGTCCAATTTCTCAGTATTGGCAACAGAGGCTGGTAAACTCAGACAAAC 300
Db 565 GCCATTTGAGAGTGTAAATTTCTCAGTACTGGCAACAGAGGAGGAGAACTCAGACAGCA 624
QY 301 ATTGACATTTTGCAAAATGCAAAATAGGCAATTTGATGGGTGACCGGCTTACAGCTTTGA 360
Db 625 ATTGATATTTGCAAAATGCAAAATAGACACTTTGATGGGAGCGGCTTACATCTTTAAGT 684
QY 361 ATTGAAGNACTCAGACACTTGAAGTTCGACTTGAAGAAAGAAATCAGCCGAGTCGATCC 420
Db 685 GTAAAGGAGCTTAAAGCAGCTAGAAAATTCGACTTCAGAGGGGCACTTAGCAGGTTTCGATCA 744
QY 421 AAAAAGAACGAGATGTTGCTTTGAAGAGATCGACATCATGCAAGAGAGGGAACACATCTT 480
Db 745 AAGAGAGATGAATGTTGCTTTGAGAGATAGAGATTATGCAAGAGAGGGAACACATATTA 804

481	QY	ATTCAGGAGAAATGAGATTCTTCGCAGCAAGATAGCCGAGTGTCAGAATAGCCACAAACG	540
805	Db	CTGGCGGAGAACACAGTTTCTTCGCACCTAAGATAGCTGAATGTGAAGCAGCGCCAAATGCA	864
541	QY	AAACATGTTATCAGCTCCGGGAAATATGATGCATCTGCCGCATTCGACTCTTCGAAATTTCCCTA	600
865	Db	AATATGCTGCCGGTCCGAAATTCGATGCATTCGCTGGATTTGATTTCTCGACATTTCCCTA	924
601	QY	CATGCAAAATCTAAATCGATCGGGCCCATCACTATGCACATCAGGAAACAAACAGCTTCAG	660
925	Db	CATGCAAGTATAATGGAAT---GCTCACCATTATGCACACACAGGATCAAACGGCCCTTCAA	981
661	QY	CTTGCTCTCAACGTT---GAACGGGTGACGCTTAAACCTCAATCAAGGCAACCCGAAA--	715
982	Db	CTTGGAATGAAGTTTATGGCGGGATATATTACTCAAGCTCAATCAAGGTTCTTGGAAATG	1041
716	QY	-----ATATGCTAGTAACTTTGAATGAGATTCCAGAGTCCAAATATTCGAGGC	763
1042	Db	TCTCTAGACTTGTTGTGCAAGACATACCACCTAAAAATTTCCGTTGCCACATATGTATGAG	1101
764	QY	AGAGCACAATGGAGAGATAGCTCCTAGTATGGAATATGGAATTTATGATATTACATATG	823
1102	Db	AAAAACCAATGGGAATTTGGTAGCTACCAAGTACGAA-----TATGATAGTAAAGCTCTG	1153
824	QY	GTTCCTCAGCTTTAAATATATAGCTGTTTGGAAACAAAGATCAACATATATAGTAGTATT	883
1154	Db	CTTTATCA-ATTTCTACATCCAGATGTGGAAATTTAAAAACCAATAT- GTTTATGATT	1211
884	QY	TTTTGGCGCATGTTATCTTTCTGTTG	909
1212	Db	TCTAGCAAAATGCTATCTTTCTATTG	1237

RESULT 9			
AF492455			
LOCUS	AF492455	1164 bp	mRNA linear
DEFINITION	Cycas edentata MADS-box transcription factor (AG)		PLN 02-JUL-2002
			mRNA, complete

```

/codec=stdmads_box_transcription_factor"
/protein_id="P4M74074.1"
/db_xref="GI_21667496"
/translation="MGKRGKIEIKTENTINRQVTFCKRNLCKKAYELVCLDAEVA
LIVFSRGRIVEPANNVRKTRERYKCAQTGGGALSSNSQYQWQKGLRQVID
LQANLNRHMGDALTSLSVKYELKGLRLEGLRSLVRKSKNMLLEITEIMORREH
LAENQPLRTKAIYEVSNQNTNVLIPGFEDLPAFDSPNSPLNHLITAAAHVYTDOD
"
/notes="cyas protein"

```

SOURCE Gnetum gnetum
ORGANISM Gnetum gnetum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Gnetophyta; Gnetales; Gnetales; Gnetum.
TITLE Winter, K.U., Becker, A., Munster, T., Kim, J.T., Saedler, H. and
JOURNAL Theissen, G.
MEDLINE MADS-box genes reveal that gnetophytes are more closely related to
PUBMED confifers than to flowering plants
93307411 Proc. Natl. Acad. Sci. U.S.A. 96 (13), 7342-7347 (1999)
10377416
REFERENCE 2 (bases 1 to 1154)
AUTHORS Theissen, G.
TITLE Direct Submission
JOURNAL Submitted (12-JAN-1999) Theissen G., Molecular Plant Genetics,
Max-Planck-Institut für Breeding Reserch, Carl-von-Linne Weg 10,
50829 Koeln, GERMANY
FEATURES
source
1..1154
/organism="Gnetum gnetum"
/mol_type="mRNA"
/sub_species="gnemon"
/db_xref="taxon:3382"
1..1154
/gene="ggm3"
88..831
/gene="ggm3"
/codon_start=1
/product="putative MADS domain transcription factor GGM3"
/protein_id="CAB44449.1"
/db_xref="GI:5019431"
/translation="MGRGKIEIKRIENTNTRQVTFCKRRNGLLKXAYELSVLCDAEVA
LIVFSRRLYEPMNSVKRTIERYKTCADNNOGGAIAESNAQYQOEAVKLKQIID
VLNNQIRHYWECLOQSWTKELKQLEKLEKLGVRSKNEKLEDITLORREDNL
IRENETRINKIACSQHQHANLITAAVEYDALPRAYSRNFPHANLEIAAAHHYA
QQQETALHGSSEKSHYHYDPDPAHKFQI"
BASE COUNT 413 a 238 c 255 g 248 t
ORIGIN
Query Match 36.1%; Score 328.6; DB 8; Length 1154;
Best Local Similarity 69.9%; Pred. No. 6.2e-79;
Matches 480; Conservative 0; Mismatches 189; Indels 18; Gaps 2;
QY 1 ATGGGTCTGTGGAGATTGAGATAAGAGATTGAAATACTACGACCGACAGGTCAT 60
DB 88 ATGGGACGCGCAAGATCGAGATAAGAGATCGAGACACACCAACACAGAGTAACT 147
QY 61 TTCTGCAAGCCCGAATGTTTATTAAGAGGCGTATGAATTATCAGTTCTTTGTGAT 120
DB 148 TTCTGCAAGCCCGAATGTCCTGCTGAAGAGGCGCTACAGATTCTCTCGCAT 207
QY 121 GCAGAGTGGCCCTCATCGTCTTCTCAGCAGAGGAGACTTTATGAATTTGCCAACAC 180
DB 208 GCTGAGTGCCTCATCGTTTCTCAGCGAGGCGTCTTATGAATTCGCCAACAC 267
QY 181 AGCTGAGAGGAGCATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 240
DB 268 AGTGTCAAGAGACGATCGAGAGTATAGAAAGACTTGTGCGGACACATCATCAGGAG 327
QY 241 GCGATATCAGATCCAAATTCCTCATGTTTGCACAGAGGCGTGGTAACTTCACACACAG 300
DB 328 GCAATTCAGAAATCCACGACAGTATTTGCCAGCAGGAGCTGTGAATTCGAGAGC 387
QY 301 ATTGACATTTTGCAAAATGCAAAATGAGCATTTGATGGTGAGCGGCTTACAGCTTTGA 360
DB 988 ATAGAGCTCTCAACACACAGATCCGACATACATGCGGAGATGCTTTCAGTCCATG 447
QY 361 ATTAAGGACTCAGACACTTGGGTCGACTTGAAGAGGATCAGCCGAGTGCATCC 420
DB 448 ATCAAGAACTGAAGACAGCTGGAGGTTAACTTCGAGAAAGACTTGTGATAGTTGATCC 507
QY 421 AAAAAGAACGAGATGTTGCTTTGAAGAGATCGACATCATGACAGAGGGAACACATCTT 480

Db 508 AAAAGGAATGAGAACTGCTGGAGCATAGACACCTTGCAGAGAGGAGATACCTG 567
QY 481 ATCCAGAGAAATGAGATTCCTGCGAGCAAGATAGCCAGTGTGAGATAGCCACACG 540
DB 568 ATTGAGAGAAATGAGTACATTCGAGAACAGATAGCAGATGCCAGAGCCATCAGATGCA 627
QY 541 AACATGTTATCAGTCC-----GGAATATGATGCACTGCCGCGATTTCGACTTCGA 591
DB 628 AACATGTTGATGCGCTGCTGTAGATGATGATGATGATGATGATGATGATGATGATG 687
QY 592 AATTTCCTACATGCAAAATCTAATGATGCGC-----CCATCACTATGACATCAG 642
DB 688 AACTTCATGATGCAAACTGATAGAGCAGCAGCTGCTCATCATCATCATCATCATCA 747
QY 643 GAACAAACACGCTTCAGCTTGCTGA 669
DB 748 GAGCAAACTGCGCTTCATCTTGGATCA 774
RESULT 11
AB050644 912 bp mRNA linear PLN 31-OCT-2001
LOCUS Magnolia praecocissima mRNA for putative MADS-domain transcription
DEFINITION factor MpMADS2, partial cds.
ACCESSION AB050644
VERSION AB050644.1 GI:16549059
KEYWORDS Magnolia praecocissima
SOURCE Magnolia praecocissima
ORGANISM Magnolia praecocissima
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; magnoliids; Magnoliales;
Magnoliaceae; Magnolia.
REFERENCE 1
AUTHORS Ito, M., Shiobara, S., Tanabe, Y. and Hasebe, M.
TITLE Organ identities in Magnolian flower
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 912)
AUTHORS Ito, M., Shiobara, S. and Hasebe, M.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-2000) Motomi Ito, University of Tokyo, Department
of Systems Sciences, Graduate School of Arts and Sciences; 3-8-1
Komaba, Meguro-ku, Tokyo 153-8902, Japan
(E-mail: mto@mail.ecc.u-tokyo.ac.jp, Tel:81-3-5454-6638,
Fax:81-3-5454-6638)
FEATURES
source
1..912
/organism="Magnolia praecocissima"
/mol_type="mRNA"
/db_xref="taxon:81865"
1..912
/gene="MpMADS2"
CDS
1..627
/gene="MpMADS2"
/codon_start=1
/product="putative MADS-domain transcription factor
MpMADS2"
/protein_id="BAB70737.1"
/db_xref="GI:16549060"
translaton="NRQVTFCKRRNGLLKXAYELSVLCDAEVALIVFSRRLYEYAN
NSVNTIDRYKACADSSGSCVSEANSYQQSSKURQIQIALLQANRHLMEALS
SMYVKELKLEKLEKISIRSKNELLFABIEYMKQREVLDQNDNNYLRAKITENE
RAQQCMGLPAPEDVMPGDFSRNFLQVNLMDSSHHYSHQEQALQSG"
912..911
polyA_signal
/gene="MpMADS2"
BASE COUNT 282 a 184 c 199 g 247 t
ORIGIN
Query Match 33.4%; Score 303.8; DB 8; Length 912;
Best Local Similarity 67.7%; Pred. No. 3.9e-72;
Matches 441; Conservative 0; Mismatches 207; Indels 3; Gaps 1;
QY 46 AACCCAGGTCACCTTCTGCAAGCGCGGAATGTTTATTAAAGAGCGGTATGAATTA 105


```

Db 1 AACAGCAAGTCACTTTTTCAGCGCGGTAATGCGCTCTCTCAAGAAGGCGCTATGAAC 60
Qy 106 TCAGTCTTTTGTGATGAGAGTGGCCCTCATCGTCTTCTCCAGCAGGAGACTTTAT 165
Db 61 TCTGTGCTTTGTGATGCTGAGTGGCTCTCATCGTCTTCTCCAGCGCGCGCTCTAC 120
Qy 166 GAATTTGCCAACACAGCGGTGAGAGGACGATTCAGAGGTACAGAGACTTCGGTTGAC 225
Db 121 GAATACGCCCAATAACAGTGTAGGAATACATCGATAGGTATAGAGAGGATGTCAGAT 180
Qy 226 AACAAACACAGGAGGCGGATATCAGAGTCCCAATCTCAGTATTCGCAACAGGAGGTGGT 285
Db 181 TCCTCAAGTTTAGGATGCGTATCGGAAGCCCAATCTCAGTATTCACAGCAAGATCATCA 240
Qy 286 AACTCAGACACACATTTGCAATTTTGCATATCAATAGGCAATTTGATGGGTGACGG 345
Db 241 AACTGCGCCCAACAAATAGCGCTTTTACAGAAATCAACAGGCAATTTGATGGGTGAGGC 300
Qy 346 CTTACAGCTTTGAACATTTAAGGAATCAAGCAACTTGAAGTTTGCATTTGAAAAGGAATC 405
Db 301 CTCAGTTTCATGCTGTTAAGGAGCTCAAGCAGCTGGAGAAATAGATTGGAGAAAGCAAT 360
Qy 406 ASCCGAGTGCATCCAAAGCAAGAGATGTTGCTGAGAGATCGACATCATCGAGAGA 465
Db 361 AGCAAGATCAGATCAAGAGAGAGAGTGTCTGTTGCTGAAATCGAGTACATCGAGAA 420
Qy 466 AGGGAACACATCTTATCCAGGAGATGAGATTTCTCGCAAGATAGCGGAGTGTCTAG 525
Db 421 AGGAGAGTTGATTTGCAAAATGATAATATGTAACCTCGTCTGCTAGAGAACGAA 480
Qy 526 AATAGC---CACACACACATGTTATCAGCTCGGAATATGATGATGCTGCGCGCATTC 582
Db 481 AGAGCCCAACAGCAATGGGATGTTGCTGCTCTCTGAATACAGATGTAATGCTGATTT 540
Qy 583 GACTCTCGAAATTTCTACATGCAATCTAATCGATCGCGCCCATCATCTATGCACATCAG 642
Db 541 GATTCCTGGAATCTCTCAAGTAATCTAATGATTCAGTCCATCATCTCTCACCAG 600
Qy 643 GACAAACACAGCTTCAGCTGGCTGACGTTGAGCGGTGGAGCTTAA 693
Db 601 GAGCAACAGCGCTTCAACTGGCTTAAAGTAATGTCGTGGATGGTCCAA 651

```

RESULT 12
AF373604
LOCUS Vitis vinifera MADS-box protein 5 (MADS5) mRNA, complete cds.
ACCESSION AF373604
VERSION AF373604.1
KEYWORDS GI:20385589
SOURCE Vitis vinifera
ORGANISM Vitis vinifera

REFERENCE
AUTHORS Bess, P.K., Sensi, E., Hua, C., Davies, C. and Thomas, M.R.
TITLE Cloning and characterization of grapevine (Vitis vinifera L.) MADS-box genes expressed during inflorescence and berry development
JOURNAL Plant Sci. 162 (6), 887-895 (2002)
AUTHORS Bess, P.K., Sensi, E., Hua, C., Davies, C. and Thomas, M.R.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2001) Plant Industry, CSIRO, Hartley Grove, Urrbrae, SA 5064, Australia

FEATURES
source
1..942
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
1..942

gene

```

CDS
/gene="MADS5"
/1..742
/gene="MADS5"
/notes="VvMADS5"
/codon_start=1
/product="MADS-box protein 5"
/protein_id="AA211345.1"
/db_xref="GI:20385590"
/translation="MGRKIBIKRIENTNRQVTFCKRRNGLKKAYELSVLCDAEVA
LIVFSRRGVYVYNNIKSTIDRYKASDSSTNGFTWENAINAQYQOESAKLRQIQ
MLQNSRHLMGDSLAVTKELQLENRLERGITRIRSKHLLAEIYLOKRETEL
ENESVYLRTKIAEVRILQANMVSIFNAIQALVSRNFPQPNMIBGSGTGYPLHDK
VLHLG"
BASE COUNT 300 a 183 c 216 g 243 t
ORIGIN
Query Match 30.8%; Score 279.6; DB 8; Length 942;
Best Local Similarity 64.3%; Pred. No. 1.7e-65;
Matches 436; Conservative 0; Mismatches 239; Indels 3; Gaps 1;
Qy 1 ATGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAAATACTACGAACCGCAGAGTCACT 60
Db 71 ATGGGGAGAGGAAGATCGAGATCAAGAGATCGAATAACAGACCAACCGTCAGTCA 130
Qy 61 TTCTGCAAGCGCCGAATGTTTATTAAGAAGCGGTATGAATATCAGTTCTTTGTGAT 120
Db 131 TTCTGCAAGCGAAGGAATGGGCTTTTGAAGAGGCTTATGAATATCAGTCTATGAT 190
Qy 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGAGCTTTATCAATTTTGCAACAC 180
Db 191 GCAGAGTGGCCCTCATCGTCTTCTCCAGCGCGGTGAGTCTATGAGTACTCAACAC 250
Qy 181 AGCGTGAAGAGGAGATTTGAGAGGTACAAGAGACTTTGGCTTGACAACACACGAGGG 240
Db 251 AACATAAATCAACATAGATAGGTACAAGAGGCGGCTCAGATAGTACAAATGGAGGC 310
Qy 241 GCGATATCAGAGTCCCAATCTCAGTATTTGGCAACAGAGGCTGGTAACTCAGACAAAC 300
Db 311 TTTACATGAGATCAATGCCAATATACCAGCAAGATCAGCAAGCTGCCACGAA 370
Qy 301 ATTGACATTTTGCAAAATGCAATAGGCAATTTGATGGGTGACGGGCTTACAGTTTGAAC 360
Db 371 ATACAGATGCTGCAGAAATCTTAACAGGCAATTAATGGGTGATTCCTTGGCTTCTT 430
Qy 361 ATTAAGGAATCAAGCAACTTGAGGTTGCACTTTGAAAAGGAATCAGCCGAGTCCGATCC 420
Db 431 GTGAGAGGATTAAGCAGCTCGAGAACAGCTTGACAGGCGATCACAGAAATCAGTTCG 490
Qy 421 AAAAAGAACAGATGTTGCTTTGAAGAGATCGACATCATCGAGAGAGGAGAACATACT 480
Db 491 AAGAAGCATGAGTTGCTGTTGGCTGAGATTGATATCTTCAGAAAAGGAAATTTGAGCT 550
Qy 481 ATCCAGAGATGAGATTTCTTCGACAGATAGCGGAGTTCAGAAATAGCCACACAG 540
Db 551 GAAATGAAGCGTATATCTCCGAAACAGATTTGCAGAGTGGAGAGGCTTCAGCAAGCA 610
Qy 541 AACATGTTTACGCTCGGAAATGATGATGCTCCGCGGATTCGACTCTCGAAATTTCTTA 600
Db 611 AACATGTTTACACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 670
Qy 601 CATGCAATCTAATTCGA---TGCGGCGCCATCCTATGTCATCAGGAAACAAACAGCT 657
Db 671 CAGCCCAATATGATGAGGCTGGATCCAGAGGCTACCCACTTCATGATGAAGAGGCTTC 730
Qy 658 CAGCTGGCTGAAGCTTG 675
Db 731 CATCTCGGTAATGATG 748

```

RESULT 13
AF234617
LOCUS Phalaenopsis equestris MADS-box transcription factor (MADS1) mRNA,
DEFINITION AF234617 1009 bp mRNA linear PLN 12-FEB-2002

```

complete cds.
ACCESSION AF234617 GI:18650788
VERSION AF234617.1
SOURCE Phalaenopsis equestris
ORGANISM Phalaenopsis equestris
Phalaenopsis equestris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
Epidendroideae; higher Epidendroideae; Vandaeae; Aeridinae;
Phalaenopsis.
REFERENCE 1 (bases 1 to 1009)
AUTHORS Chen,H.H., Liu,C.C., Tsuei,S.W. and Chen,W.H.
TITLE Phalaenopsis equestris MADS-box protein (MADS1)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1009)
AUTHORS Chen,H.H. and Tsuei,S.W.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Institute of Biotechnology, National Cheng
Kung University, Tainan, Taiwan 701
REFERENCE 3 (bases 1 to 1009)
AUTHORS Chen,W.H.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Department of Horticulture, Taiwan Sugar
Research Institute, Tainan 702, Taiwan
REFERENCE 4 (bases 1 to 1009)
AUTHORS Liu,C.C.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Department of Biology, National Cheng Kung
University, Tainan 701, Taiwan
FEATURES
source
1..1009
Location/Qualifiers
/organism="Phalaenopsis equestris"
/mol_type="mRNA"
/db_xref="taxon:78828"
/clone="pMADS1"
/tissue_type="flower bud"
1..1009
/gene="MADS1"
/notes="homeotic gene"
112..831
/gene="MADS1"
/codon_start=1
/product="MADS-box transcription factor"
/protein_id="AAL76415.1"
/db_xref="GI:18650789"
/translation="MDSSSEPEKMGKIBIKIENTNRQVTFCKRNLGLKAY
ELSVLDAVALIIIFSTRGLRYEYANNVSKVITRIYKASTDNTSISANSQYQ
QATKLROQITNLQNSRLNLDGTLMSLRDLKOLETRLEKINKIRAKNELHAE
IDYMKREVELCTDNMFLRNKISDNERAQQHMSILPSTSTEVEMPPDPSRSLF
VNLMDPDRYSHQQTALQLG"
BASE COUNT 337 a 188 c 235 g 249 t
ORIGIN
Query Match 30.7%; Score 279; DB 8; Length 1009;
Best Local Similarity 65.4%; Pred. No. 2.5e-65;
Matches 449; Conservative 0; Mismatches 220; Indels 18; Gaps 2;
QY 1 ATGGGTCGTGGGAAGATTGAGATAAGAGATTGAAATCTACGACCGACGAGTCACT 60
DDB 145 ATGGGAGGGGGAGATAGAGATCAAGAGAATAGACACACACAAACAGATTACC 204
QY 61 TTCTGCAAGCGCGGAATGTTTATTAAAGAGCGGTATGAAATATCAGTTCTTTGTGAT 120
DDB 205 TTTTTCAGCGCGGTATGCGCTCTCTGAGAGAGCTTATGAGCTCTCTGTTGTGAT 264
QY 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAATTCGCCAACAC 180
DDB 265 GCTGAGGTTGCCCTAATCATCTTCTACCCGTGGCGGCTCTATGAATATGCAACAC 324
QY 181 AGCGTGNAGAGGAGCTTGAGAGGTACAGAGACTTGGTTGACACACACACGAGGG 240
DDB 325 AGCGTGAAGGAGACCATTTGAACGATACAAAGAGGAGCACTGATTAATTCCTGGA 384

```

```

241 GCGATATCAGAGTCCAAATTCAGTATTGGCAACAGAGGCTGGTAAACTCAGACACAG 300
385 TCATATCTGAGCAATTCGAGTATTACCAACAGAGAGGCCACAAAACCTCGCTCAGCA 444
301 ATTGACATTTTGCAAAATGCAATAGCATTGATGGGTGACGGGCTTACAGCTTTGAAC 360
445 ATTACAAACTTACAGAAATTCACAGGAATTTGCTGGGTGATGCTCTCACCACCATGAGC 504
361 ATTAAGCAACTCAGCAACTTGGAGTTTCGACTTTGAAAAGGAATCAGCCGAGTGCATCC 420
505 CTGAGGACCTGAGCAACTTGGAAAACAGATTGGAGAAAGGCATCACAAAATAAGAGCT 564
421 AAAAAGAACGAGATGTTTGGTTTGAAGAGATCGACATCATCGAGAGAGGGAACACATCTT 480
565 AAGAAGAAATGAATCTGCTGCTGATGCTGAGATTGACTACATGCGAGAAAAGGAAATGGAATC 624
481 ATCCAGAGAGATGAGATCTTTCGCGCAAGATGCGCA-----GTGTGAGATATGCCACAA 536
625 CAAACTGACACACATGTTTCTGCGCAATAGATATCTGATTAATGAAGAGCAGACAGCAG 684
537 CACGAACATGTTATCAGCT-----CCGGAATATGATGACATGCGCCGCAATC 582
685 CATCAGCATATGAGCATATTGCCATCAACAGCAGACAGATATGAAGTATGCTCTCCATTT 744
583 GACTCTCGAAATTTCTACATGCAATCTAATCTAATGATCGCGGCCCATCACTATGACATCAG 642
745 GATTCCAGAAAGCTTTCTTCATGTCATCTAATGATCCCAATGACCGTTATTCACCACAG 804
643 GAACAAACAAAGCTTTCAGTTTGGCTGA 669
805 CAGCAAAACAGCTCTGCAACTTGGGTGA 831

RESULT 14
AF022377 1142 bp mRNA linear PLN 17-AUG-2001
LOCUS Cucumis sativus agamous-like putative transcription factor (CAG1)
DEFINITION mRNA, complete cds.
ACCESSION AF022377
VERSION AF022377.1 GI:4103341
SOURCE Cucumis sativus (cucumber)
ORGANISM Cucumis sativus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
REFERENCE 1 (bases 1 to 1142)
AUTHORS Perl-Treves,R., Kahana,A., Rosenman,N., Xiang,Y. and Silberstein,L.
TITLE Expression of multiple AGAMOUS-like genes in male and female
flowers of cucumber (Cucumis sativus L.)
JOURNAL Plant Cell Physiol. 39 (7), 701-710 (1998)
MEDLINE 98399593
PUBMED 9729894
REFERENCE 2 (bases 1 to 1142)
AUTHORS Perl-Treves,R., Rosenman,N., Kahana,A. and Silberstein,L.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-1997) Life Science, Bar-Ilan University,
Ramat-Gan 52900, Israel
FEATURES
source
1..1142
Location/Qualifiers
/organism="Cucumis sativus"
/mol_type="mRNA"
/db_xref="taxon:3659"
/tissue_type="flower"
/dev_stage="bud"
1..1142
/gene="CAG1"
173..850
/gene="CAG1"
/codon_start=1
/product="agamous-like putative transcription factor"
/protein_id="AAD01742.1"
/db_xref="GI:4103342"

```

/translation="MGRGKIEIKRIENTNRQVTFCKRRNGLLKKAYELSVLCDAEVA
LIVFSRRRLVYSNNISIKTTIRYKACSDSSATSVTELNTQYVQBSAKLRQIO
MLONSNRLMGDSLSALTVELKOLENLERGIPIRSKHEMLAEIYLOKRETEL
ENENYCIJTKIAEVERYQANWVSQELNATQALANSRFFSPNMEGAVTFSHQD
KQWHLHG"

BASE COUNT 413 a 193 c 209 g 327 t

ORIGIN
Query Match 30.5%; Score 276.8; DB 8; Length 1142;
Best Local Similarity 65.2%; Pred. No. 1e-64;
Matches 442; Conservative 0; Mismatches 227; Indels 9; Gaps 2;

QY 1 ATGGGTCTGTGGGAAGATTGAGATAAGAGGATTGAAATATACTACGACCGACAGGTCCT 60
DB 173 ATGGGGAGAGGAAAGATAGATTAAGAGATAGAGAACACAAATAGCAAGTTACA 232
QY 61 TTCTCGAAGCGCCGAATGGTTTATTAAAGAGCGGTATGAATATATCAGTCTCTTTGTGAT 120
DB 233 TTCTGTAAGAGAGAAATGGACATTTTGAAGAAAGCTTATGAATCTCTGTTCTTTGTGAT 292
QY 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGAGAGGAGACTTTATGAATTTGCCAACCCAC 180
DB 293 GCTGAAGTTGCTCTCATTTGTTCTCCAGCGGTGCGGCTCTCTATGATTAATCAATAAC 352
QY 181 AGCGTGAAGAGACGAATTGAGAGGTACAAAGAGACTTCGTTGACACACACACGAGGG 240
DB 353 AGCATCAAAACAACTATTGAGAGGTACAAAGAGGCTTGTCTGATAGCTCAGCTACTAGC 412
QY 241 GCGATATCAGAGTCCAAATTCCTCAGTATTGGCAACAGGAGGCTGGTAACTCAGACAAAC 300
DB 413 TCTGTCACTGAACATAACTCAATATATCAGCAGAGATCGGCTAACTGCGTCAACAG 472
QY 301 ATTGACATTTTGCAAAATGCAATAGGCAATAGGCAATTTGATGGTGACGGGCTTACAGTTTGAAC 360
DB 473 ATACAAATGCTTCAGAAATTCACAGGCACTTGATGGGGGACTCCTTGAGTGTCTTACT 532
QY 361 ATTAAGGAACCTCAAGCAACTTGGAGTTCGACTTGAAGAAAGAAATCAGCCGAGTGCATCC 420
DB 533 GTCAGAACTCAGCAGCTTGAATAGGCTTGAAGAGGATCACTAGATCAGATCA 592
QY 421 AAAAGAGACGAGATGTTCTTGAAGAGATGCAATCATGACAGAGAGGGAACACATATT 480
DB 593 AAGAAGACGAAATGTTCTAGCAGAAATTCAGTACCTTCAGAAAAGGAGATTGAGCTG 652
QY 481 ATCCAGGAGATGAGATCTTTCGAGCAAGATAGCCAGTGTGAGATAGCCACACACG 540
DB 653 GAGACGAAATGTGTATTAGAACCAAGATAGCAGAGTAGAGGGTTCAACAGCA 712
QY 541 AACATGTTATCAGCTCCGGAATATGATGCTAGTCCGCGCATTCG ---ACTCTGAAATTC 597
DB 713 AACATGTTATCGGACCAAGAACTGAATGCAATACAGCAATTGGCTAACTCTGCAATTC 772
QY 598 CTACATGCAAAATCTAATCGA-----TGCGGCCCATCATCATGCAATCAGGAAACAAACA 651
DB 773 TTCTCTCCCAATATCATGGAACCTGCTGGACCTGTTCTTACTCTCATCAAGCAAGAA 832
QY 652 AGCTTTCAGTTGGGTGA 669
DB 833 ATGCTTTCATCTGGGTGA 850

RESULT 15

AY178837

LOCUS

DEFINITION

Memordica charantia mads-box transcription factor mRNA, complete

cds.

ACCESSION

AY178837

VERSION

AY178837.1

KEYWORDS

Memordica charantia (balsam pear)

ORGANISM

Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE

Yang, M., Lan, L., Zeng, Y., Zhang, M., Bai, J., Miao, C., Cai, Y.,

Lai, J., Xu, X., and Chen, F.

Direct Submission

JOURNAL

Submitted (12-NOV-2002) Sichuan University, #29 Wangjiang Road,

Chengdu, Sichuan 610064, P. R. China

Location/Qualifiers

1. 1001

/organism="Memordica charantia"

/mol_type="mRNA"

/db_xref="taxon:3673"

51..734

/codon_start=1

/product="mads-box transcription factor"

/protein_id="AAC20104.1"

/db_xref="GI:27763670"

/translation="MGRGKIEIKRIENTNRQVTFCKRRNGLLKKAYELSVLCDAEVA

LIVFSRRRLVYSNNISIKTTIRYKACSDSSATSVTELNTQYVQBSAKLRQIO

MLONSNRLMGDSLSALTVELKOLENLERGIPIRSKHEMLAEIYLOKRETEL

ENENYCIJTKIAEVERYQANWVSQELNATQALANSRFFSPNMEGAVTFSHQD

KQWHLHG"

QDKMLHLG"

BASE COUNT 337 a 162 c 239 g 263 t

ORIGIN

Query Match 29.3%; Score 266; DB 8; Length 1001;

Best Local Similarity 64.2%; Pred. No. 9e-62;

Matches 439; Conservative 0; Mismatches 230; Indels 15; Gaps 2;

QY 1 ATGGGTCTGTGGGAAGATTGAGATAAGAGGATTGAAATATACTACGACCGACAGGTCCT 60

DB 51 ATGGGGAGAGGAAAGATCGAGATTAAAGCGGATCGAGAACACGACAAATCGGAGGTGAG 110

QY 61 TTCTGCAAGCGCCGAATGGTTTATTAAAGAGAGGCGGTATGAATATATCAGTCTCTTTGTGAT 120

DB 111 TTCTGCAAGAGGAGGAATGACCTGCTGGAAGAAAGCTTATGAATCTGCTCTTTGTGAT 170

QY 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACAC 180

DB 171 GCTGAAGTTCCTCATCGTCTTCTCCAGCGGTGCGGCTCTATGAATCTATCCAAATAC 230

QY 181 AGCGTGAAGAGGACGATTGAGAGGTACAAAGAGACTTCGCTTGACAAACACACGAGGG 240

DB 231 AGCATCAAAATACCATTCGGAGGTACAAAGAGGCTTGTCTGATAGCTCAGCTACTAGC 290

QY 241 GCGATATCAGAGTCCAAATTCAGTATTGGCAACAGGAGGCTGTTAACTCAGACAAAC 300

DB 291 TCTGTCACTGAATCAATACTCAATATATCAGCAGAAATCTGCAAGCTGCGCAACAA 350

QY 301 ATTGACATTTTGCAAAATGCAAA-----TAGGCATTTGATGGGTGACGGGCTT 348

DB 351 ATACAAATGCTACAGAAATCCACAGCAATCTTGTAGGCACTTGTATGGGGACTCCTTG 410

QY 349 ACAGCTTTGAACTTAAGGAATCAAGCAATTCAGGTTGAGTTGAAAGAAAGAAATCAGC 408

DB 411 AGTGCTCTTACAGTGAAGGAATTTGAAGAGCTTGAAGATAGGCTTGAAGAGGATCACT 470

QY 409 CGAGTGCATCCAAAAGAACGAGATGTTGCTTTGAAGAGATCGACATCATGACAGAGAGG 468

DB 471 AGGATCAGATCAAGAGGATGAAATGTTGCTAGCAGAAATTTAGTACTCTGACAGAAAGA 530

QY 469 GAACACATCTTATCCAGGAGATGAGATTTCTTCGACGAGATAGCCGAGTGTGAGAT 528

DB 531 GAGATCAGCTGGAGATGAAATGTTGTTAGAACCAAGATAGCAGAGTGGAGAGG 590

QY 529 AGCCACAAACAGAACTGTTATCAGCTCCGGAATATGATGCACTCCCGCATTCGACTCT 588

DB 591 CTTTCAGCAAGCCAAACATGGTTTCTGGGCAAGAACTGAATGCAATTCAGGCATTTGCTTCT 650

QY 589 CGAAATTTCTTACATGCAATCTAATCGA---TGCGGCCCATCATCATGCAATCAGAA 645

DB 651 CGAAATTTCTTACTCTTAAATGATGAGAGGTGAGCTGTCACCTTCTCACCACCAAGAC 710

QY 646 CAACACACCTTCAGCTTGGCTGA 669
Db | | | | | | | | | | | | | | | |
711 AAGAGATGCTTCATCTTGGTGA 734

Search completed: January 21, 2004, 22:14:26
Job time : 3322.05 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 15:50:10 ; Search time 234.397 Seconds
(without alignments)
10468.497 Million cell updates/sec

Title: US-09-936-869-3

Perfect score: 909

Sequence: 1 atgggtcgtgggaagattga.....cgcattgtatctttctgttg 909

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	909	100.0	909	21	AAA99364
2	260	28.6	1201	21	AAZ57947
3	260	28.6	1219	22	AAF85398
4	260	28.6	1219	24	ABK84889
5	259.8	28.6	723	22	AAF85399
6	255.6	28.1	777	21	AAZ57947
7	255.6	28.1	959	20	AAV99860
8	255.6	28.1	959	24	ABQ79456

9	255.6	28.1	959	25	ABX11511
10	251.4	27.7	1157	21	AAZ50201
11	251.4	27.7	1157	21	AAZ35494
12	248.2	27.3	714	22	AAF85402
13	248.2	27.3	1115	21	AAZ57949
14	248.2	27.3	1159	22	AAF85401
15	248.2	27.3	1159	24	ABK88491
16	248.2	27.3	1297	24	ABK88491
17	245.2	27.0	706	21	AAZ55501
18	245.2	27.0	706	21	AAZ55501
19	245	27.0	896	20	AAV99859
20	245	27.0	896	24	ABQ79455
21	245	27.0	896	25	ABX11510
22	236.2	26.0	1202	21	AAZ34634
23	235.6	25.9	903	21	AAZ34432
24	232.8	25.6	1143	17	AAZ34428
25	230.2	25.3	1057	21	AAZ34428
26	230.2	25.3	1057	21	AAZ34428
27	229.2	25.2	805	24	ABK82084
28	229.2	25.2	1259	24	ABK82084
29	217.8	24.0	1218	21	AAZ34634
30	209.8	23.1	581	21	AAZ34634
31	200.2	22.0	1321	21	AAZ34634
32	175.6	19.3	1144	19	AAV18014
33	175	19.3	557	21	AAZ34634
34	175	19.3	1070	19	AAV18012
35	172.8	19.0	466	21	AAZ34634
36	172.4	19.0	385	21	AAZ34634
37	172.4	19.0	385	21	AAZ34634
38	171.2	18.8	1268	22	AAZ34634
39	171.2	18.8	1268	22	AAZ34634
40	171.2	18.8	1280	24	AAZ34634
41	167.2	18.4	511	21	AAZ34634
42	165.6	18.2	558	21	AAZ34634
43	164.6	18.1	687	21	AAZ34634
44	164.6	18.1	1038	24	ABK85164
45	164.6	18.1	1038	25	AAZ34634

ALIGNMENTS

RESULT 1

AAA99364

ID AAA99364 standard; DNA; 909 BP.

AC AAA99364;

XX 22-JAN-2001 (first entry)

XX Plant PrAG1 promoter DNA sequence.

DE Plant promoter; PrAG1; reproductive tissue; transgenic plant; cereal; ds.

XX Plant promoter; PrAG1; reproductive tissue; transgenic plant; cereal; ds.

XX Pinus radiata.

OS WO200055172-A1.

XX WO200055172-A1.

XX 21-SEP-2006.

XX 17-MAR-2000; 2000WO-NZ00031.

XX 17-MAR-1999; 99NZ-0334715.

XX (CART-) CARTER HOLT HARVEY LTD.

PA (TASM-) TASMAN BIOTECHNOLOGY LTD.

PA (UNMT-) UNIV MICHIGAN TECHNOLOGICAL.

XX Podila GK, Liu J, Karnosky DF;

XX MPI, 2000-524442/56.

XX Novel plant reproductive tissue promoter, useful to produce plants

DNA encoding Arabi
petunia hybrida ne
petunia nectary-sp
Nucleotide sequenc
Poplar floral home
Nucleotide sequenc
Poplar floral home
Sweetgum cDNA enco
Eucalyptus grandis
Arabidopsis AGU1 c
Arabidopsis AGU1 e
DNA encoding Arabi
Arabidopsis thalia
Eucalyptus AGE2 CD
Arabidopsis thalia
Eucalyptus AGE1 CD
Arabidopsis thalia
Arabidopsis thalia
DNA encoding novel
DNA encoding novel
Zea mays DNA fragm
Eucalyptus grandis
Zea mays DNA fragm
Pinus radiata cone
Pinus radiata tran
Pinus radiata tran
Zea mays DNA fragm
Pinus radiata tran
Maize MADS-box gen
Maize ZmMADS2 codi
Corn nitrate-respo
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia

PT which have a diminished reproductive capacity or which are sterile -
XX Claim 8; Page 39-40; 51pp; English.
XX This invention relates to a novel plant promoter gene. The promoter is
CC located in plant reproductive tissue, and the invention includes
CC transgenic plants containing the promoter. The promoter can be used to
CC produce plants which have a diminished reproductive capacity or which are
CC sterile. The constructs can also be used to transform agronomically
CC important plants in which modulation of reproductive capacity
CC (particularly the timing and abundance of flowering) is desirable,
CC e.g. cereals, rice, maize, wheat, barley, oats, rye, soybean and canola.
CC The present sequence represents DNA encoding the plant reproductive
CC promoter (PTAG1) of the invention.
XX
XX Sequence 909 BP; 299 A; 172 C; 216 G; 222 T; 0 other;

Query Match 100.0%; Score 909; DB 21; Length 909;
Best Local Similarity 100.0%; Pred. No. 3.4e-261;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGTCTGGGAAGATTGAGATAAAGAGGATTGAAATATCTACGACCCGACAGGTCACT 60
DB 1 ATGGGTCTGGGAAGATTGAGATAAAGAGGATTGAAATATCTACGACCCGACAGGTCACT 60
QY 61 TTCTCAAGCCGCGAATGGTTTATTAAGAGCGGTATGATATATCAGTCTTTGTGAT 120
DB 61 TTCTCAAGCCGCGAATGGTTTATTAAGAGCGGTATGATATATCAGTCTTTGTGAT 120
QY 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACCCAC 180
DB 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACCCAC 180
QY 181 AGCGTGAAGAGGAGGATTGAGAGGTACAGAGAGTTCGCTTGACACACACCGAGGG 240
DB 181 AGCGTGAAGAGGAGGATTGAGAGGTACAGAGAGTTCGCTTGACACACACCGAGGG 240
QY 241 GCGATATCAGATGCTCAATCTTCAGTATGCGAAGAGGCTGTTAACTCAGACACACAG 300
DB 241 GCGATATCAGATGCTCAATCTTCAGTATGCGAAGAGGCTGTTAACTCAGACACACAG 300
QY 301 ATTGACATTTGCAAAATGCAATAGGATTTGATGGTGACGGCTTACAGCTTTGAAC 360
DB 301 ATTGACATTTGCAAAATGCAATAGGATTTGATGGTGACGGCTTACAGCTTTGAAC 360
QY 361 ATTAAGGAACCTCAAGCACTTTGAGGTTGCTGCTTCAAAAGGAATTCAGCCGAGTGCATCC 420
DB 361 ATTAAGGAACCTCAAGCACTTTGAGGTTGCTGCTTCAAAAGGAATTCAGCCGAGTGCATCC 420
QY 421 AAAAGACGAGATGTTCTTGAAGAGATCGACATCTGACAGAGAGGAGACATACCTT 480
DB 421 AAAAGACGAGATGTTCTTGAAGAGATCGACATCTGACAGAGAGGAGACATACCTT 480
QY 481 ATCCAGGAATGAGATTTCTTCGAGCAAGATAGCCGAGTGTGAGATAGCCCAACACG 540
DB 481 ATCCAGGAATGAGATTTCTTCGAGCAAGATAGCCGAGTGTGAGATAGCCCAACACG 540
QY 541 AACATGTTATCAGCTCCGGATATGATGCTACCTGCGCGCATTCGACTCTGAAATTTCCCTA 600
DB 541 AACATGTTATCAGCTCCGGATATGATGCTACCTGCGCGCATTCGACTCTGAAATTTCCCTA 600
QY 601 CATGCAAAATCTAATCGATGCGGCCCATCATCTATCATCAGGACCAACACAGCTTCAG 660
DB 601 CATGCAAAATCTAATCGATGCGGCCCATCATCTATCATCAGGACCAACACAGCTTCAG 660
QY 661 CTGCTGAAAGCTTGAAGCGGTGACGCTTAAACTCAATCAAGGACCCGAAATATG 720
DB 661 CTGCTGAAAGCTTGAAGCGGTGACGCTTAAACTCAATCAAGGACCCGAAATATG 720
QY 721 CTAGTAACTTGAATGAGATTGAGATTCAGAGTCTGAAATATTCGAGGCAAGACCAATGGAAGA 780
DB 721 CTAGTAACTTGAATGAGATTGAGATTCAGAGTCTGAAATATTCGAGGCAAGACCAATGGAAGA 780

QY 781 GATAGCTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GATAGCTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 ATAGCTCTTTGAAACAAAGAAATACAAATATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 ATAGCTCTTTGAAACAAAGAAATACAAATATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 TTTCTGTTG 909
DB 901 TTTCTGTTG 909
RESULT 2
AAZ57947
ID AAZ57947 standard; cDNA; 1201 BP.
XX AC AAZ57947;
XX DT 11-APR-2000 (first entry)
XX DE Poplar floral homeotic gene PTAG-1 cDNA.
XX KW Poplar; PTAG-1; floral homeotic gene; transgenic plant; sterility;
XX KW fertility; ss.
XX OS Populus balsamifera subsp. trichocarpa.
XX FH Key Location/Qualifiers
XX FT CDS 196..921
XX FT /*tag= a
XX FT /note= "the coding region is also specifically
XX PN CA2227940-A1.
XX PD 06-OCT-1999.
XX PF 07-APR-1998; 98CA-2227940.
XX PR 06-APR-1998; 98US-0080851.
XX PA (UYOR-) UNIV OREGON HEALTH SCI.
XX PI Rottmann WH, Brunner AM, Sheppard LA, Strauss SH;
XX DR WPI; 2000-105662/10.
XX P-PSDB; AAY58656.
XX PT Nucleic acid from Populus trichocarpa genes, useful for producing
XX PT transgenic plants, particularly trees, with modified fertility
XX PT characteristics such as sterility -
XX PS Claim-1; Page 64-66, 32pp; English.
XX The present sequence is that of cDNA encoding Populus balsamifera
XX subsp. trichocarpa PTAG-1 protein (see AAY58656). The PTAG-1 gene (see
XX AAZ57946) is 1 of 4 newly identified floral homeotic genes from this
XX poplar species. PTAG-1 is a homologue of AGAMOUS and is expressed
XX in floral tissue. The invention provides nucleic acid sequences of
XX these 4 Populus genes, the corresponding cDNA sequences (see
XX AAZ57942-49) and deduced amino acid sequences (see AAY58454-57). It also
XX provides methods of using the genes and cDNA to produce genetically
XX engineered Populus species and other trees having modified fertility
XX characteristics, including sterility. Genetic constructs useful in
XX producing these genetically engineered trees include antisense
XX versions of PTAG-1, dominant negative mutants, and constructs useful
XX for sense suppression. Promoter sequences may be used to obtain
XX floral specific expression of genes such as cytotoxins that may be
XX used in genetic ablation strategies to produce trees having modified
XX fertility characteristics, including sterility. Sterile trees
XX allow increased wood yield and a reduction in the production of
XX allergens such as pollen.

```

XX SQ Sequence 1201 BP; 354 A; 251 C; 257 G; 339 T; 0 other;
Query Match      28.6%; Score 260; DB 21; Length 1201;
Best Local Similarity 63.7%; Pred. No. 3.8e-67;
Matches 433; Conservative 0; Mismatches 235; Indels 12; Gaps 2;

QY 2 TGGTCTGCGGAGATTGAGATAAGAGAGATTGAAATACTACGAACCGACAGGTCACTT 61
DB 242 TGGGAGGGGAAAGGTGGAGATCAAGCGGATCGAGAACACCAACCAATGCCAAGTCACTT 301

QY 62 TCTGCAAGCGCGGAATGTTTAAAGAACGCGTATGAATTATCAGTTCTTTCGTGATG 121
DB 302 TCTGCAAAAGGCGCATGTTGCTCAAGAAAGCCCTACGAATATCTGTTCTTTCGTGATG 361

QY 122 CAGAAGTGGCCCTCAFCGTCTTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACCA 181
DB 362 CTGAGGTTCACATCATCGTCTTCTTAGCGCGGTGCGCTTTATGAGTACTCTAACGATA 421

QY 182 GCGTGAAGAGGAGGATTGAGGATCAAGAAAGACTTGCCTTACAAACACGAGGAGGG 241
DB 422 GTGTCAAAATCAAAATGAGAGGTACAAAAGGCAATCTGCAGATTTCTTCAAAACACTGGGT 481

QY 242 CGATATCAGAGTCCAAATCTCAGTATTGGCAACAGAGGCTGGTAACTCAGACCAACAGA 301
DB 482 CTGTTTCTGAGCCATGCTCAGTACTACCAGCAAGAGCTGCCAAGCTGCCGTCCCAAA 541

QY 302 TTGACATTTTGCMAATGCAATAGGCATTTGATGGGTGACGGCTTACAGCTTTGAACA 361
DB 542 TTGGTAAATTTGCAGAAATTCAAACAGGCATATGCTGGGTGAAGCGCTTATGTTCAITGAGTG 601

QY 362 TTAAGAACTCAAGCAACTTGAGTTTCGACTTTGAAAGAAAGAAATCAGCGAGTGCATCCA 421
DB 602 TGAAGAACTTAAAGTTTGAATACGACTTTGAGAAAGAAATAGCAGAAATTCGTTCCA 661

QY 422 AAAAGAACGAGATGTTGCTTGAAGAGATCGACATATGTCAGAGAGAGGGAACACATAC 481
DB 662 AAAAGAAATGAGTGTGTTTGCAGAAATCCAGTATATGTCAGAGAGAGGAGGTTGACTTGC 721

QY 482 TCAGAGAGATGAGATTTCTTCGACGACAGATAGCGGAGTGCA---GAATAGCCACACA 538
DB 722 ACACAAATACACAGCTTCTCCGAGCAAGAAATTTTACAGAAATGAAGAGAGCAGACAGA 781

QY 539 CGAATATGTTATCAGCTCCGGAATATGATGCACT-----GCCCGCATTCGACTCTC 589
DB 782 TGAATTTGATGTCAGGAGGAGCAGACTTTTGAATCGTGCACTCTCAACCATATGACTCTC 841

QY 590 GAAATTTCTACATGCAATCTAAATCGATGCGGCCCATCACTATGCACATCAGGACAAA 649
DB 842 GGAATCTATCTCAAGTGAATGGAATACAGCCTGCAAGTCACTACTACATCAAGATCAGA 901

QY 650 CAACGCTTCAGCTTGCTGA 669
DB 902 TGSCCCTTCAGTTAGTTTAA 921

RESULT 3
ID AAF85398
XX AAF85398 standard; cDNA; 1219 BP.
XX AC AAF85398;
XX XX
DT 23-JUL-2001 (first entry)
XX
XX Nucleotide sequence of the floral homeotic protein PTAG-1.
DE PTAG-1; PTAG-1; PTAG-2; floral tissue; LEAFY;
KW LEAFY; FLORICAULA; FLO; DEFICIENS; DEF; AGAMOUS; AG; transgenic plant;
KW fertility; sterility; ss.
XX
XX Populus balsamifera.
OS
XX
XX Key Location/Qualifiers

```

```

CDS
FT 196..921
FT /*tag= a
FT /product= "PTAG-1"
CA2319853-A1.
XX 01-APR-2001.
XX 02-OCT-2000; 2000CA-2319853.
XX 01-OCT-1999; 99US-0410464.
XX (UYOR-) UNIV OREGON HEALTH SCI.
XX Rottman WH, Straus SH, Brunner AM, Sheppard LA;
XX WPI; 2001-336098/36.
XX P-PSDB; AAB68437.
XX Novel isolated polynucleotide derived from Populus species, useful for
XX producing transgenic plants having modified fertility characteristics,
XX particularly sterility -
XX Claim 25; Page 54-55; 69pp; English.
XX The present sequence encodes a floral homeotic protein, designated
XX PTAG-1, which is derived from Populus balsamifera subsp. trichocarpa.
XX The specification also describes PTD, PTLF, and PTAG-2 proteins. The
XX floral homeotic proteins are expressed in floral tissues. PTLF is a
XX homologue of LEAFY (LFY) and FLORICAULA (FLO), and is expressed in
XX immature inflorescences on which floral primordia are developing. PTD
XX is a homologue of DEFICIENS (DEF), and is strongly expressed in stamen
XX primordia from the onset of organogenesis. PTAG-1 and PTAG-2 are
XX homologues of AGAMOUS (AG). The floral homeotic proteins and
XX polynucleotides are useful for producing transgenic plants having
XX modified fertility characteristics, particularly sterility.
XX SQ Sequence 1219 BP; 371 A; 252 C; 257 G; 339 T; 0 other;

Query Match      28.6%; Score 260; DB 22; Length 1219;
Best Local Similarity 63.7%; Pred. No. 3.8e-67;
Matches 433; Conservative 0; Mismatches 235; Indels 12; Gaps 2;

QY 2 TGGTCTGCGGAGATTGAGATAAGAGAGATTGAAATACTACGAACCGACAGGTCACTT 61
DB 242 TGGGAGGGGAAAGGTGGAGATCAAGCGGATCGAGAACACCAACCAATGCCAAGTCACTT 301

QY 62 TCTGCAAGCGCGGAATGTTTAAAGAAAGCGTATGAATATCACTTCTTGTGATG 121
DB 302 TCTGCAAAAGGCGCATGTTGCTCAAGAAAGCCCTACGAATATCTGTTCTTTCGTGATG 361

QY 122 CAGAAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACCA 181
DB 362 CTGAGGTTCACATCATCGTCTTCTTAGCGCGGTGCGCTTTATGAGTACTCTAACGATA 421

QY 182 GCGTGAAGAGGAGGATTGAGGATGAGAGAGACTTGCCTTGAACCAACACGAGGAGGG 241
DB 422 GTGTCAAAATCAAAATGAGAGGTACAAAAGGCAATCTGCAGATTTCTTCAACACTGGGT 481

QY 242 CGATATCAGAGTCCAAATTTCTCAGTATTGGCAACAGGAGGCTGGTAACTCAGACCA 301
DB 482 CTGTTTCTGAGCCATGCTCAGTACTACCAGCAAGAGCTGCCAAGCTGCCGTCCCAAA 541

QY 302 TTGACATTTTGCMAATGCAATAGGCATTTGATGGGTGACGGCTTACAGCTTTGAACA 361
DB 542 TTGGTAAATTTGCAGAAATTCAAACAGGCATATGCTGGGTGAAGCGCTTATGTTCAITGAGTG 601

QY 362 TTAAGAACTCAAGCAACTTGAGTTTCGACTTTGAAAGAAAGAAATCAGCGAGTGCATCCA 421
DB 602 TGAAGAACTTAAAGTTTGAATACGACTTTGAGAAAGAAATAGCAGAAATTCGTTCCA 661

QY 422 AAAAGAACGAGATGTTGCTTGAAGAGATCGACATATGTCAGAGAGAGGGAACACATAC 481
DB 662 AAAAGAAATGAGTGTGTTTGCAGAAATCCAGTATATGTCAGAGAGAGGAGGTTGACTTGC 721

QY 482 TCAGAGAGATGAGATTTCTTCGACGACAGATAGCGGAGTGCA---GAATAGCCACACA 538
DB 722 ACACAAATACACAGCTTCTCCGAGCAAGAAATTTTACAGAAATGAAGAGAGCAGACAGA 781

QY 539 CGAATATGTTATCAGCTCCGGAATATGATGCACT-----GCCCGCATTCGACTCTC 589
DB 782 TGAATTTGATGTCAGGAGGAGCAGACTTTTGAATCGTGCACTCTCAACCATATGACTCTC 841

QY 590 GAAATTTCTACATGCAATCTAAATCGATGCGGCCCATCACTATGCACATCAGGACAAA 649
DB 842 GGAATCTATCTCAAGTGAATGGAATACAGCCTGCAAGTCACTACTACATCAAGATCAGA 901

QY 650 CAACGCTTCAGCTTGCTGA 669
DB 902 TGSCCCTTCAGTTAGTTTAA 921

RESULT 3
ID AAF85398
XX AAF85398 standard; cDNA; 1219 BP.
XX AC AAF85398;
XX XX
DT 23-JUL-2001 (first entry)
XX
XX Nucleotide sequence of the floral homeotic protein PTAG-1.
DE PTAG-1; PTAG-1; PTAG-2; floral tissue; LEAFY;
KW LEAFY; FLORICAULA; FLO; DEFICIENS; DEF; AGAMOUS; AG; transgenic plant;
KW fertility; sterility; ss.
XX
XX Populus balsamifera.
OS
XX
XX Key Location/Qualifiers

```

Db 662 AAAAGAATGAGCTGTTGTTGTCAGAAATCGAGTATATGACAGAGAGGAGGTTGACTTGC 721
 QY 482 TCAGGAGAAATGAGATTTCTTCGACGAAATAGCGAGTGTC---GAATAGCCACAACA 538
 Db 722 ACAACAATAACAGCTTCTCCGAGCAAGAATTTGAGAGAAATGAAGAAGCGACAGACA 781
 QY 539 CGAATGTTATCACTCGGAAATATGATGCACT-----GCCCGATTGACTCTC 589
 Db 782 TGAATTTGATGCGACGAGAGAGACACTTTTGAGATCGTGCGAGTCTCAACCATATGACTCTC 841
 QY 590 GAAATTTCTATCAATCAATCTTAATCGATGCGGCGCCATCACTATGACATCAGCAACA 649
 Db 842 GGAATCTTCTCAAGTGAATGATTAACAGCTGCAAGTCACTACTACATCAAGATCAGA 901
 QY 650 CAACGCTTCAGCTTGCTGA 669
 Db 902 TGGCCCTTCAGTTAGTTTAA 921

RESULT 4
 ABK8489
 ID ABK8489 standard; cDNA; 1219 BP.
 AC ABK8489;
 XX
 XX 07-OCT-2002 (first entry)
 DT
 DE Poplar floral homeotic protein, PTAG-1, cDNA.
 KW Poplar; ss; gene; plant; DEFICIENS; transgenic; promoter;
 KW protein transduction domain; floral homeotic gene;
 KW floral-specific expression; cytotoxin; fertility; sterility;
 KW PTLF; PTAG-1; PTAG-2.
 OS Populus balsamifera subsp. trichocarpa.
 XX
 XX Key Location/Qualifiers
 FH 196..921
 CDS /tag= a
 FT /product= "PTAG-1"
 FT
 FT
 XX US9359892-B1.
 XX
 XX 28-MAY-2002.
 XX
 XX 01-OCT-1999; 99US-0410464.
 XX
 XX 06-APR-1998; 98US-080851P.
 PR 06-APR-1999; 99US-0287700.
 XX
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 XX Strauss SH, Rottmann W, Brunner A, Sheppard L;
 PI
 XX WPI; 2002-572853/61.
 DR P-PSDB; AB330867.
 DR
 XX New protein transduction domain promoter nucleic acid molecule useful
 PT for producing transgenic plants having modified fertility
 PT characteristics, particularly sterility -
 XX
 XX Disclosure; Column 63-66; 46pp; English.
 XX
 XX The invention relates to an isolated nucleic acid molecule especially a
 CC protein transduction domain (PTD) promoter; (i) that hybridises under
 CC wash conditions of 0.2 x SSC (saline sodium citrate), 0.1 %SDS (sodium
 CC dodecyl sulphate) at 65 plusoc to nucleotides or (ii) comprising 35
 CC consecutive nucleotides of the PTD gene. PTD is a floral homeotic
 CC gene and is the homologue of DEFICIENS. Also includes are a recombinant
 CC nucleic acid comprising the PTD promoter, a cell transformed with
 CC the recombinant nucleic acid and a transgenic plant comprising the
 CC transformed cell. The PTD promoter is useful to obtain floral-specific
 CC expression of genes such as cytotoxins, that are employed in genetic

CC ablation strategies to produce trees having modified fertility
 CC characteristics, including sterility. Genetic constructs comprising
 CC antisense versions or dominant negative mutants of PTD are useful in
 CC producing genetically engineered Poplars and other trees, and for sense
 CC suppression. Also disclosed are 3 other homeotic genes PTLF, PTAG-1
 CC and PTAG-2 (none are defined). The present sequence is the PTAG-1 cDNA.
 XX

SQ Sequence 1219 BP; 371 A; 252 C; 257 G; 339 T; 0 other;

Query Match 28.6%; Score 260; DB 24; Length 1219;
 Best Local Similarity 63.7%; Pred. No. 3.8e-67;
 Matches 433; Conservative 0; Mismatches 235; Indels 12; Gaps 2;

QY 2 TGGCTCGTGGGAACATTCAGATTAAGAGGATTCGAATATCTAGAACCGACAGTCACATT 61
 Db 242 TGGGAGGGGAAAGTGGAGATCAAGCGGATCGAAGAACCCCAATGCGCAAGTCATT 301
 QY 62 TCTCAAGCGCCGAAATGGTTTATTAAAGAGCGGTATGAATTTATCAGTTCTTTTGATG 121
 Db 302 TCTGCAAAAGCGCAGTGGTTTGTCTCAAGAAAGCCTACGAATTTATCTGTTCTTGGCATG 361
 QY 122 CAGAAAGTGGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTTATGAATTTGCCAACACA 181
 Db 362 CTGAGTTGCACTCATCTCTTCTTAGCGCGGTTCCTTTATGAGTACTCTAAGATA 421
 QY 182 GCGTGAAGAGGACGATTGAGAGGTACAAGAAAGCTTCGTTGACAAACACCGAGGGGG 241
 Db 422 GTGTCAAAATCAACAATTCAGAGGTACAAAAGGCATCTGCAGATTTCTTCAAACTGGGT 481
 QY 242 CGATATCAGAGTCCCAATTCCTCAGTATTGGCAACAGAGGCTGTTAACTCAGACAACA 301
 Db 482 CTGTTTCTGAAGCAATGCTCAGTACTACCAGCAAGAGCTGCCAGCTCGTTCCTCCAA 541
 QY 302 TTGACATTTTGCAAAATGCAAAATAGGATTTGATGGGTGACGGCTTACAGCTTTGAACA 361
 Db 542 TTGTTAATTTGCAGATTTCAACAGGCATATGCTGGGTGAAGCGCTTAGTTTCATTGAGTG 601
 QY 362 TTAAGGAACCTCAGCAACTTTGAGTTTGAATTTCAAAAAGAAATCAGCGAGTGCATCCA 421
 Db 602 TGAAGGAACCTTAAGATTTGGAATACGACTTGAAGAAAGGATTAAGCAATTCGTTCCA 661
 QY 422 AAAGAAACGAGATGTTCTTTGAAGAGATCGACATCATGTCAGAGAGAGGGGAACATCTTA 481
 Db 662 AAAGAAATGAGCTGTTGTTGCGAGAAATCGAGTATATGTCAGAGAGAGGGGTTGACTTC 721
 QY 482 TCCAGGAGAAATGAGATTTCTTCGAGCAAGATAGCCGAGTGTC---GAATAGCCACAACA 538
 Db 722 ACAACAATACCAGCTTCTCCGAGCAAGATTTTCAGAGATGAAGAAGCGACAGACA 781
 QY 539 CGAACATGTTATCAGTCCCGAATATGATGCACT-----GCCCGATTGACTCTC 589
 Db 782 TGAATTTGATGCCAGGAGGAGCAGACTTTGAGATCGTGAGTCTCAACCATATGACTCTC 841
 QY 590 GAAATTTCTTACATGCAAAATCTTAATCGATGCGGCGCCATCACTATGACATCAGGAACAAA 649
 Db 842 GGAATATTCTCAAGTGAATGGATTACAGCTCGAAGTCACTACTACATCAGATCAGA 901
 QY 650 CAACGCTTCAGTTGGCTGA 669
 Db 902 TGGCCCTTCAGTTAGTTTAA 921

RESULT 5

AAF85399
 ID AAF85399 standard; cDNA; 723 BP.

XX AAF85399;
 AC AAF85399;

DT 23-JUL-2001 (first entry)

DE Nucleotide sequence of the floral homeotic protein PTAG-1.

XX Floral homeotic gene; PTD; PTLF; PTAG-1; PTAG-2; floral tissue; LEAFY;
 KW

KW LEY; FLORICAULA; FLO; DEFICIENS; DBF; AGAMOUS; AG; transgenic plant;
 KW fertility; sterility; ss.
 XX
 OS Populus balsamifera.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..723
 FT /*tag= a
 FT /product= "PTAG-1"
 XX
 XX CA2319853-A1.
 XX
 XX 01-APR-2001.
 XX
 XX 02-OCT-2000; 2000CA-2319853.
 XX
 XX 01-OCT-1999; 99US-0410464.
 XX
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 XX Rottman WH, Strauss SH, Brunner AM, Sheppard LA;
 XX
 XX WPI; 2001-336098/36.
 XX P-PSDB; AAB68437.
 XX
 XX Novel isolated polynucleotide derived from Populus species, useful for
 PT producing transgenic plants having modified fertility characteristic,
 PT particularly sterility -
 XX
 XX Claim 25; Page 55-56; 69pp; English.
 XX
 XX The present sequence encodes a floral homeotic protein, designated
 CC PTAG-1, which is derived from Populus balsamifera subsp. trichocarpa.
 CC The specification also describes PTD, PTF, and PTAG-2 proteins. The
 CC floral homeotic proteins are expressed in floral tissues. PTF is a
 CC homologue of LFAY1 (LFY) and FLORICAULA (FLO), and is expressed in
 CC immature inflorescences on which floral primordia are developing. PTD
 CC is a homologue of DEFICIENS (DBF), and is strongly expressed in stamen
 CC primordia from the onset of organogenesis. PTAG-1 and PTAG-2 are
 CC homologues of AGAMOUS (AG). The floral homeotic proteins and
 CC polynucleotides are useful for producing transgenic plants having
 CC modified fertility characteristics, particularly sterility.
 XX
 XX Sequence 723 BP; 226 A; 147 C; 175 G; 175 T; 0 other;
 SQ

Query Match 28.6%; Score 259.8; DB 22; Length 723;
 Best Local Similarity 63.9%; Pred. No. 3.4e-67;
 Matches 431; Conservative 0; Mismatches 232; Indels 12; Gaps 2;
 QY 2 TGGTCTGTTGGAGATTGAGATAAAGAGGATTGAAAATCTAGCAACCGACAGGTTCACATT 61
 DB 47 TGGGAGGGGAAGGTGGAGATCAGCGGATCAGAACACCAATGCCAAGTCACATT 106
 QY 62 TCTGCAAGCCCGAATGGTTTATTAAGAGCGGTATGAAATATCAGTCTCTTTGGATG 121
 DB 107 TCTGCAAAAGCCGAGGTGTTGCTCAAGAAAGCCCTACGAAATATCTGTTCTTTGGATG 166
 QY 122 CAGAAGTGGCCCTCATCTCTTCCAGCAGAGGAGACTTTATGAATTTGCCAACCA 181
 DB 167 CTGAGTTGCACTCATCTCTTCTTAGCCGCGGTGCTTTATGAGTACTCTAACGATA 226
 QY 182 GCSTGAAGAGGACGATTGAGAGGTACAAAGAACTTCGCTTGACAAACACCGAGGGGG 241
 DB 227 GTGTCAAAATCAAAATGAGAGGTACAAAAGGCTCTGCAGATTCTTCAAAACACTGGGT 286
 QY 242 CGATATCAGATCCAAATCTTCAGTTATGGCAACAGGAGGCTGTAACTCAGACACACA 301
 DB 287 CTGTTTCTGAAGCAATGCTCAGTACTACAGCAAGAGCTGCCAAGCTGCGTTCCTCAA 346
 QY 302 TTGACATTTTGCAAAATGCAAAATAGGCATTGTGAGGTGACGGGCTTACAGCTTTGAACA 361
 DB 347 TTGTAATTTGCAGAAATTCAAACAGGCAATATGCTGGGTGAAGGCTTAGTTCAATTGAGTG 406

QY 362 TTAAGGAACCTCAGCAAACTTGAGTTGAGTTGAAATGAAAGGAATCAGCCGAGTGGATCCA 421
 DB 407 TGAAGGAACCTTAAGAGTTTGAATATACGACTTGAGAAAGGAATAAGCAGAAATTCGTTCCA 466
 QY 422 AAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATGCGAGAGGGAACACATACTTA 481
 DB 467 AAAAGAAATGAGCTGTTGTTTGCAGAAATCGAGTATATGCGAAGAGGAGGTTGACTTGC 526
 QY 482 TCCAGGAGAAATGAGATTCTTCGCAAGCAAGATAGCCGAGTGTCA---GAATAGCCACAACA 538
 DB 527 ACAACAATACCAGCTTCTCCGAGCAAGATTTCAGAGAAATGAAGAAAGGACGACAGCA 586
 QY 539 CGAACATGTTATCAGCTCCGGAATATGATGACT-----GCCGCAATTCGACTTCTC 589
 DB 587 TGAATTTGATGCCAGGAGGACGACTTTGAGATCGTGCAGTCTCAACCATATGACTTCTC 646
 QY 590 GAAATTTCTACATGCAAAATCTAATCGATCGGCGCCATCATGTCACATCAGCAACAA 649
 DB 647 GGAATTTCTCAAGTGAATGGATTACAGCTGCAAGTCAATCTACATCAAGATCAGA 706
 QY 650 CAACGCTTCAGCTTG 664
 DB 707 TGGCCCTTCAGTTAG 721

RESULT 6
 AAC43464
 ID AAC43464 standard; DNA; 777 BP.
 XX
 AC AAC43464;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 39336.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 XX EPI033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 PF
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.

PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	20-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138647.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	24-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	28-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0143277.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144684.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	26-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.

```

PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      28.1%; Score 255.6; DB 21; Length 777;
Best Local Similarity 68.3%; Pred. No. 6.3e-66;
Matches 354; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 1 ATGGTCTGTGGAGAGATTGAGATAAAGAGAGATTGAAATATCTACGAAACCGACAGGTCAC 60
DB 46 ATAGGAGAGGAGATAGATAGATAGAGAGATAGAGAGATAGAGAGATAGAGATAGAGATAG 105
QY 61 TTCTCAAGCGCGGAAATGGTTTATTAAGAGAGCGGTATGAAATATATCAGTTCTTTGTGAT 120
DB 106 TTCTCAAGCGCGGAAATGGTTTATTAAGAGAGCGGTATGAAATATATCAGTTCTTTGTGAT 165
QY 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAAATTTGCCAACAC 180
DB 166 GCTGAGGTGTCTTGTTCATCTTCTCCACTCGAGGCGGTCTTACGAGTACGCCAACAC 225
QY 181 AGCGTGAAGAGAGCAGATTGAGAGGTACAGAGAGCTTTCGTTGACACACACCGAGGG 240
DB 226 AGTGTGAGAGAGAAACATAGAGAGGTACAGAGAGCTTTCGTTGACACACACCGAGGG 285
QY 241 CGGATATCAGATGCAATCTTCATATTGCGAACAGAGGCTGGTAAACTCAGACACACAG 300
DB 286 ACCATCACCGAAGCTAATCTACTACTATCAGTATCAGAGAGGCGTCTAAACTCCGGAGAC 345
QY 301 ATTGACATTTTGAAGATGCAATAGAGCTTTCGTTGAGAGGCTTACAGCTTTTGAAC 360
DB 346 ATTCGGGACATTCAGATTTCAGACAGACATCTTGTGAAATCTTTGTTCCTTTGAAC 405
QY 361 ATTAAGAGAACTCAGACAACTTGAGTTGCACTTGAAGAAAGGATCAGCCGAGTGCATCC 420
DB 406 TTTAAGGAACCTCAAGAACCTTGAAGTAGGCTTGAGAGAGGAATCAGTGTGTCCGATCC 465
QY 421 AAAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATGACAGAGAGGGAACACATATT 480
DB 466 AAGAGACGAGATGTTAGTTGACAGATTGATATCATGCAAAAAGGGAATTCGAGCTG 525
QY 481 ATCCAGGAGATGAGATTTCTTCGAGCAGCAGATAGCCGA 518
DB 526 CAAACGATACATGATCTCCGCTCCAGAGATTACTGA 563

```

RESULT 7

```

ID AAV99860 standard; cDNA; 959 BP.
XX AC AAV99860;
XX DT 10-MAY-1999 (first entry)
XX DE Arabidopsis AGL5 cDNA.
XX KW AGL5-like gene; agamous-like 5 gene; seed dispersal; dehiscence;
XX KW transgenic plant; promoter; ss.
XX OS Arabidopsis thaliana.
XX FH Key Location/Qualifiers
XX CDS 78..818
XX FT /*tag= a
XX PN W09900502-A1.
XX PD 07-JAN-1999.
XX

```

```

PF 25-JUN-1998; 98WO-US13208.
XX
PR 28-APR-1998; 99US-0067800.
PR 27-JUN-1997; 97US-0051030.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Ferrandiz C, Yanofsky MF;
XX
DR WPI, 1999-095747/08.
DR P-PSDB; AAW81000.
XX
PT Use of agamous-like nucleic acids - useful for the production of
PT transgenic seed plants in which dehiscence is modified resulting in
PT delayed seed dispersal
XX
PS Disclosure; Page 90-91; 126pp; English.
XX
CC This is the nucleotide sequence of an Arabidopsis agamous-like 5
CC (AGL5) gene cDNA clone that codes for AGL5 polypeptide (see
CC AAW81000) which is involved in the regulation of dehiscence. The
CC invention provides a transgenic seed plant, such as an agl5 and
CC agl1 double mutant, that is characterised by delayed seed dispersal
CC due to suppression of AGL5 and AGL1 (see AAV9859) expression. The
CC invention also provides dehiscence zone regulatory elements, derived
CC from the AGL1 or AGL5 gene (see AAV9857 and AAV9858), that confer
CC selective expression upon an operatively linked nucleic acid
CC molecule in the valve margin or dehiscence zone of a seed plant.
CC The transgenic seed plants include members of the Brassicaceae,
CC such as rapeseed, and members of the Fabaceae, such as soybean,
CC pea, lentil and bean. A plant expression vector comprising a
CC dehiscence zone-selective regulatory element, and a kit for
CC producing the transgenic seed plants are also provided.
XX
SQ Sequence 959 BP; 297 A; 218 C; 214 G; 230 T; 0 other;

Query Match      28.1%; Score 255.6; DB 20; Length 959;
Best Local Similarity 68.3%; Pred. No. 7e-66;
Matches 354; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 1 ATGGTCTGTGGAGAGATTGAGATAAAGAGAGATTGAAATATCTACGAAACCGACAGTCACT 50
DB 123 ATAGGAGAGGAGAGATAGATAGATAGAGAGATAGAGAGATAGAGAGATAGAGAGATAG 182
QY 61 TTCTCAAGCGCGGAAATGGTTTATTAAGAGAGCGGTATGAAATATATCAGTTCTTTGTGAT 120
DB 183 TTCTCAAGCGCGGAAATGGTTTATTAAGAGAGCGGTATGAAATATATCAGTTCTTTGTGAT 242
QY 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAAATTTGCCAACAC 180
DB 243 GCTGAGGTGTCTTGTTCATCTTCTCCACTCGAGGCGGTCTTACGAGTACGCCAACAC 302
QY 181 AGCGTGAAGAGAGCAGATTGAGAGGTACAGAGAGCTTTCGTTGACACACACCGAGGG 240
DB 303 AGTGTGAGAGAGAAACATAGAGAGGTACAGAGAGCTTTCGTTGACACACCGAGGG 362
QY 241 GCGATATCAGAGTCCCAATCTCAGTATTCGACACAGGAGGCTGTAACTCAGACACAG 300
DB 363 ACCATCACCGAAGCTAATCTACTACTATCAGTATCAGAGAGGCGTCTAACTCCGGAGAC 422
QY 301 ATTGACATTTTGAAGATGCAATAGGCAATTTGATGGGTGACGGCTTACAGCTTTTGAAC 360
DB 423 ATTCGGGACATTCAGAAATTTGAACAGACATCTTGTGTAATCTCTTGGTTCTTTGAAC 482
QY 361 ATTAAGGAACCTCAGCAACTTCAGGTTTCGACTTGAAGAGGATCAGCCGAGTGCATCC 420
DB 483 TTTAAGGAACCTCAGCAACTTCGAGTATTCGAGTATTCGAGTATTCGAGTATTCGAGTAT 542
QY 421 AAAAAGAACGAGATGTTGCTTGAAGAGATTCGACATCATGACAGAGAGGGAACACATATT 480
DB 543 AAGAGACGAGATGTTAGTTGACAGAGATTGAAATATGCAAAAAGGGAATTCGAGCTG 602
QY 481 ATCCAGGAGATGAGATTCTTCGAGCAGAGATAGCCGA 518

```

603 CAAACGATACATGTATCTCCGCTCCAGATTACTGA 640

RESULT 8

ABQ79456
ID ABQ79456 standard; DNA; 959 BP.

AC ABQ79456;

DT 15-NOV-2002 (first entry)

DE Arabidopsis AGL5 encoding sequence.

KW Plant; lignin; AGL5; lignification; vascular plant; woody plant;
KW eucalyptus; cottonwood; alder; Douglas fir; hemlock; pine; spruce;
KW leguminous plant; alfalfa; clover; lucerne; birdsfoot trefoil;
KW stylosanthes; lotononis bainesii; sainfoin; forage grass; bahiagrass;
KW bermudagrass; dallisgrass; pangolagrass; big bluestem; Indiangrass;
KW switchgrass; smooth brome grass; orchardgrass; timothy;
KW Kentucky bluegrass; tall fescue; agriculture; wood product; AGL5; gene;
ds.

OS Arabidopsis thaliana.

XX Key Location/Qualifiers
FH CDS 78..818
FT /*tag= a
FT /product= "AGL5"

XX US6410826-B1.

XX 25-JUN-2002.

XX 25-JUN-1999; 99US-0339998.

XX 25-JUN-1998; 98US-090649P.

XX (REGC) UNIV CALIFORNIA.

XX Yanofsky MF, Liljegen S, Ferrandiz C;

XX WPI; 2002-588897/63.

XX P-PSDB; ABB98194.

XX Reducing lignification in vascular plants e.g. woody plant, forage
XX grass, or leguminous plant comprises ectopically expressing a nucleic
XX acid molecule encoding an Arabidopsis AGL5-like gene product in the
XX vascular plant

XX Disclosure; Column 41-44; 4ipp; English.

XX The invention relates to reducing lignification in a vascular plant by
XX introducing an exogenous nucleic acid molecule encoding an AGL5-like
XX gene product (a MADS box protein). This results in lignification being
XX reduced due to ectopic expression of the nucleic acid molecule. The
XX method of the invention is useful for reducing lignification in a
XX vascular plant, including a woody plant, such as Eucalyptus, cottonwood,
XX alder, Douglas fir, Hemlock, pine and spruce, a leguminous plant such as
XX alfalfa, clover, lucerne, birdsfoot trefoil, stylosanthes, lotononis
XX bainesii and sainfoin, or a forage grass such as bahiagrass,
XX bermudagrass, dallisgrass, pangolagrass, big bluestem, Indiangrass,
XX switchgrass, smooth brome grass, orchardgrass, timothy, Kentucky bluegrass
XX and tall fescue. The method is useful for developing plant varieties
XX with improved characteristics. This is useful in agriculture and the wood
XX products industry, both benefiting by a reduction in lignin content. The
XX current sequence represents the Arabidopsis AGL5 encoding sequence.

SQ Sequence 959 BP; 297 A; 218 C; 214 G; 230 T; 0 other;

Query Match 28.1%; Score 255.6; D9 24; Length 959;

Best Local Similarity 68.3%; Pred. No. 76-66; Indels 0; Gaps 0;

Matches 354; Conservative 0; Mismatches 164;

QY 1 ATGGGTCTGGAGAGATTGAGATAAGAGGATTGAAATACTACGAACCGACAGGTCACT 60
Db 123 ATAGGGAGAGGGAAGATAGAGATAAGAGGATAGAGAACACTACGAATCGTCAAGTCACT 182
QY 61 TTCTGCAAGCGCCGAATGTTTATTAAAGAAGCGGTATGAATTATCAGTTCTTTTGTGAT 120
Db 183 TTCTGCAAAACGACGCAATGTTTACTCAGAAAGCTTATGAGCTCTCTGTTGTGTGAC 242
QY 121 GCAGAAATGGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTATGAAATTTGCCAACCCAC 180
Db 243 GCTGAGGTGTCTTGTCTATCTTCTCCACTCGAGGCCGTCTCTACGAGTAGCGCAACAAC 302
QY 181 ACGGTGAAGAGAGCGAATGAGAGGTACAAGAAGACTTGGCTTGACACAAACCAACCGAGGG 240
Db 303 AGTGTGAGAGGAACAATAGAAAGGTACAAGAAGCTTGTCCGAGCGCGTTAAACCTCCG 362
QY 241 GCGATATCAGAGTCCAATTTCTCAGTATTTGCAACAGAGAGCTGGTAAACTCAGACACAG 300
Db 363 ACCATCACCGAAGCTAATACTACTCAGTACTATCAGCAAGAGCGCTTAAACTCCCGAGACAG 422
QY 301 ATTGACATTTTGCAAAATGCAATAGGCATTTGATGGTTCAGCGGCTTACAGCTTTTGAAC 360
Db 423 ATTGGGACATTCAGAAATTTGAACAGACACATTTCTTGTTGATCTCTTGTCTCTTGAAC 482
QY 361 ATTAAGGAATCAAGCAACTTTGAGGTTGCACTTTGAAAAGAAATCAGCCGAGTCCGATCC 420
Db 483 TTTAAGGAATCAAGAACTTTGAAAGTAGGCTTGAGAAAGAAATCAGTCTGTGTCGATCC 542
QY 421 AAAAAGAACGAGATGTTGCTTTGAAGAGATCGACATCATCGAGAGAGGGAACACATACCTT 480
Db 543 AAGAAGCACGAGATGTTAGTTGAGAGATTTGAATACATGCAAAAAGGGAATCGAGCTG 602
QY 481 ATCCAGGAGATGAGATTTCTTCGAGCAAGATAGCCGA 518
Db 603 CAAAACGATAACATGTATCTCCGCTCCAGATTACTGA 640

RESULT 9

ABX11511

ID ABX11511 standard; DNA; 959 BP.

AC ABX11511;

DT 01-MAY-2003 (first entry)

DE DNA encoding Arabidopsis Agamous-like protein 5 (AGL5).

XX Agamous-like protein; AGL1; AGL1-like protein; vascular plant;
XX woody plant; leguminous plant; forage grass; lignification; tall fescue;
XX eucalyptus; bahiagrass; bermudagrass; dallisgrass; smooth brome grass;
XX pangolagrass; big bluestem; Indiangrass; switchgrass; alfalfa;
XX orchardgrass; timothy grass; Kentucky bluegrass; animal husbandry;
XX gene; ds; transgenic.

OS Arabidopsis sp.

XX Key Location/Qualifiers

FH CDS 78..818

FT /*tag= a

FT /product= "AGL5"

XX /note= "Agamous-like protein 5"

PN US2002129403-A1.

XX 12-SEP-2002.

XX 15-OCT-2001; 2001US-0978730.

XX 25-JUN-1998; 98US-090649P.

XX 25-JUN-1999; 99US-0339998.

XX (REGC) UNIV CALIFORNIA.

XX Yanofsky MF, Liljegren S;
XX WPI; 2003-147404/14.
DR P-PSDB; ABG75903.
XX Vascular plant lignification reduction method for genetic engineering,
PT involves expressing nucleic acid molecule encoding specific type of
PT gene product in vascular plant, ectopically
XX Disclosure; Fig 8; 46pp; English.
XX The invention describes a nucleic acid molecule encoding an AGL8-like
CC gene product comprising a amino acid sequences of an AGL8 orthologue
CC such as Arabidopsis AGL8 expressed ectopically in a vascular plant such
CC as woody plant, leguminous plant or forage grass. The nucleic acids are
CC useful for reducing lignification in woody plants such as eucalyptus,
CC leguminous forage crops such as alfalfa, and in forage grasses such as
CC bahiagrass, bermudagrass, dalliagrass, pangolagrass, big bluestem,
CC indiangrass, switchgrass, smooth bromegrass, orchardgrass, timothy,
CC Kentucky bluegrass or tall fescue used for plant/agricultural genetic
CC engineering. Modification of lignification allows more efficient use of
CC plant biomass in animal husbandry, by enabling production of
CC lignification enhanced/reduced vascular plants, and enables lignification
CC changes in monocotyledonous or dicotyledonous angiosperm or with a
CC gymnosperm, which is a seed-bearing plant with seeds not enclosed in
CC ovary. This sequence encodes Arabidopsis Agamous-like protein 5 (AGL5),
CC a protein that regulates the process of lignification.
XX
SQ Sequence 959 BP; 297 A; 218 C; 214 G; 230 T; 0 other;

Query Match 28.1%; Score 255.6; DB 25; Length 959;
Best Local Similarity 68.3%; Pred. No. 7e-66;
Matches 354; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 1 ATGGGTCTGGAGATTGAGATAAGAGGATTGAAATATCTACGACCCACAGGTCCT 60
Db 123 ATAGGAGAGAGAGATAGATTAAGAGATAGAGACACTAGATCTCAAGTCCT 192
QY 61 TTCTGCAAGCCGCAATGTTTATTAAGAGAGCGTATGAATATCAGTTCTTTGTGAT 120
Db 183 TTCTGCAAGCCGCAATGTTTATTAAGAGAGCGTATGAATATCAGTTCTTTGTGAT 242
QY 121 GCAGAGTGGCCCTCATGCTTCTCCAGCAGAGGAGACTTTATGAATTTGCCACCCAC 180
Db 243 GCTGAGGTGCTCTGTGATCTTCTCCAGCAGGCGCTCTCTACGAGTACGCCAACAC 302
QY 181 AGCGTGAAGAGGAGGATTGAGAGGTACAGAGACTTTCGTTGACAAACACACCGAGGG 240
Db 303 AGTGTGAGAGAAACAATAGAGAGGTACAGAGAGCTTTCGAGCGGTTAACCTCCG 362
QY 241 CGATATCAGAGTCCATTTCTAGTATTTGGCAAGAGGCTGTAACTCAGACACACAG 300
Db 363 ACCATCACCGAAGCTAATCTACTCAGTATATCAGAGAGGCGTCTAACTCCGAGACAG 422
QY 301 ATTGACATTTGCAAAATGCAATAGGCAATTTGATGGTGACGGCTTACAGTTTGAAC 360
Db 423 ATTGGGACATTCAGATTTGACAGACACATTTCTGTGATCTCTTGGTTCTTGAAC 482
QY 361 ATTAAGGAACCTCAAGCAATTTGAGTTTCGATTTGAAAGGAATCAGCGAGTCCGATCC 420
Db 483 TTTAAGGAACCTCAAGCAATTTGAGTTTCGATTTGAAAGGAATCAGCTGTGTCGATCC 542
QY 421 AAAAGACGAGATGCTTCTTGAAGATCGACATCATGACAGAGAGGGAACACATCTT 480
Db 543 AAGAGACGAGATGCTTCTTGAAGATCGACATCATGACAGAGAGGGAACACATCTT 602
QY 481 ATCCAGGAGATGAGATCTTTCGACAGCAAGATAGCCGA 518
Db 603 CAAACGATACATGATCTCTCGCTCCAGATTACTGA 640

RESULT 10

AAZ50201
ID AAZ50201 standard; cDNA; 1157 BP.
AC
XX AAZ50201;
XX
DT 17-MAY-2000 (first entry)
XX
DE Petunia hybrida nectary-specific protein FBP15 cDNA.
XX
KW FBP15; nectary-specific protein; metabolite; recombinant protein;
KW transgenic plant; nectar; honey; insect; bee; pharmaceutical; enzyme;
KW biotest; antioxidant; food additive; MADS-box protein; ss.
XX
OS Petunia hybrida.
XX
FH Key Location/Qualifiers
CDS 179..844
FT /*tag= a
FT /product= "FBP15 protein"
FT mat_peptide 206..361
FT /*tag= b
FT /product= "MADS-box"
FT /note= "homologous to a similar region in the Arabidopsis
FT agamous protein and Antirrhinum deficiens protein"
FT 503..616
FT /*tag= c
FT /label= "K-box"
XX
FN WO200004176-A1.
XX
XX 27-JAN-2000.
XX
XX 15-JUL-1999; 99WO-NL00453.
XX
XX 16-JUL-1999; 98EP-0202375.
PR 14-DEC-1998; 98EP-0204215.
XX
XX (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
XX
XX Creemers J, Angenent GC, Kater MM;
XX
XX WPI; 2000-182438/16.
DR P-PSDB; AAY44804.
XX
PT New DNA encoding Petunia hybrida nectary-specific proteins, useful for,
e.g. producing modified honey
XX
XX Claim 12; Page 43; 93pp; English.
XX
XX The present sequence is a cDNA encoding nectary-specific protein FBP15.
CC It was isolated from a cDNA library made from nectaries of Petunia
CC hybrida strain W15 flowers. FBP15 is a MADS-box protein which is
CC exclusively expressed in the nectaries of P. hybrida.
CC A DNA sequence from the promoter region upstream of
CC nectary-specific expressed sequence e.g. NECL and FBP15 DNAs is used in
CC a recombinant DNA construct comprising a DNA encoding a metabolite
CC preferably recombinant protein, a DNA encoding a signal peptide that
CC targets the recombinant protein to the nectar and optionally a signal
CC sequence function in plants for the transcription termination and
CC polyadenylation of an RNA molecule. The DNA construct is useful for
CC producing transgenic plants which excrete recombinant proteins in its
CC nectar. The nectar is processed into honey by insects (preferably bees)
CC and the desired protein is easily recovered from it. The recombinant
CC proteins are useful for pharmaceutical purposes, as enzymes for biotests
CC and antioxidants for food additives.
XX
SQ Sequence 1157 BP; 391 A; 203 C; 229 G; 334 T; 0 other;

Query Match 27.7%; Score 251.4; DB 21; Length 1157;
Best Local Similarity 62.0%; Pred. No. 1.4e-64;
Matches 415; Conservative 0; Mismatches 251; Indels 3; Gaps 1;

QY 1 ATGGGTCTGGAGATTGAGATAAGAGGATTGAAATATCTACGACCCACAGGTCCT 60

179 ATGGGAAGAGGAAGATTGAGATTAGAGGATTGAAATACAAATCGTCAAGTCACT 238
 61 TTCTGCAAGCGCCGAATGTTTATTAAGAAGAGGCGTATGAATATCAGTTCTTTGTGAT 120
 239 TTCTGTAAGAGAGAAATGGTTGCTTAAAGAGCTTATGAATTTCTGTTCTTTGTGAT 298
 121 GCAGAGTGGCTCTGCTCTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACAC 180
 299 GCTGAAGTTGCTCTCATGCTTTCTCAAGCGGTGGCGCTCTATGAATATGCTAACAAC 358
 181 AGCGTGAAGAGGAGATTGAGAGGTACAAGAGACTTGGTTTGACAAACACCGAGGG 240
 359 AGTGTGAAGCAACAATGATAGATATAAGAAAGCATCTCAGATTCTCCACACTGGA 418
 241 GCATATCAGAGTCCCAATCTCAGTATTGGCAACAGGAGGCTGTGAATCTCAGAACAG 300
 419 TCTACTTCTGAAGCTAACACTCAGTTTATCAACAGAGAGCTGCCAACTCCGAGTTGAG 478
 301 ATTGACATTTGCAAAATGCAATAGGCAATTTGATGGGTGACGGCTTTACAGTTTGAAC 360
 479 ATTGTAACCTTACAGAACTCAACAGGAACATGCTAGCGAGTCTCTAAGTTCTCTGACT 538
 361 ATTAAGGAACCTCAAGCAACTTGAGGTTGACTTTGAAAGGAATCAGCGGAGTGGATCC 420
 539 GCAAAAGATCTGAAGGCTGGAGACCAACTTGAAGGAATTTAGTAGAATTAGGTCC 598
 421 AAAAAAGAGAGATGTTGCTTTGAAGAGATCGACATCATGAGAGAGAGGGAACATACCTT 480
 599 AAAAAAGATGAACTCTCTGTTGCTTAAAGAGCTTATGAATTTCTGTTCTTTGTGAT 658
 481 ATCCAGGAGATGAGATTTCTCGCAGCAGATAGCGGAGTGTGAGATAGCCACACAG 540
 659 CACAACACATCAGATGCTTGGGCAAGATAGCTGAGAGTGAAGAAATGTGAACATG 718
 541 AACATGTTATCAGCTCCGGAATATGATGCACTGCGCGCATTCGACTCTCGAAATTTCTTA 600
 719 ---ATGGAGGAGAAATTTGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 775
 601 CATGCAATCTAATCGATGGCGCCATCATCTATGACATCAGGAACAAACAGCTTTCAG 660
 776 CAAGTGAAGCGCTTACAGCAATATCATCAATATCCAGCCCAAGACACATGGCTCTTCAA 835
 661 CTTGGCTGA 669
 836 TTAGTATAA 844

RESULT 11

ID AAZ35494 standard; cDNA; 1157 BP.
 AC AAZ35494;
 DT 11-APR-2000 (first entry);
 DE Petunia nectary-specific FBP15 cDNA.
 XX
 XX FBP14; nectary; nectar; transgenic plant; honey; ds.
 XX Petunia hybrida.
 XX Key Location/Qualifiers
 XX CDS 179..841
 XX /*tag= a
 XX
 XX EP974667-A1.
 XX
 XX 26-JAN-2000.
 XX
 XX 16-JUL-1998; 98EP-0202375.
 XX
 XX 16-JUL-1998; 98EP-0202375.
 PR

XX PA (CPRO-) CPRO-DLO CENT PLANTENVERDELINGS REPROD.
 XX PI Creemers J, Angenent GC, Kater MW;
 XX WPI: 2000-108400/10.
 XX DR P-PSDB; AAY58448.
 XX Novel DNA sequences used to produce modified honey, the metabolites of
 XX PT which can be isolated and purified -
 XX PS Claim 5; Page 17; 56pp; English.
 XX CC The present sequence is that of Petunia hybrida strain W15 FBP15
 CC cDNA that was identified in nectary gland cDNA by differential
 CC expression. The cDNA encodes a 221-amino acid MADS box protein
 CC (see AAY58448) that is specifically expressed in the nectaries of
 CC Petunia. The present invention provides a method for producing
 CC recombinant proteins in honey. The honey is manufactured by insects,
 CC preferably honeybees, that collect the nectar of transgenic plants.
 CC The FBP15 gene and its promoter can be utilised in expression
 CC cassettes for the production of transgenic plants that produce a
 CC protein of interest in their nectar.
 XX SQ Sequence 1157 BP; 391 A; 203 C; 229 G; 334 T; 0 other;
 Query Match 27.7%; Score 251.4; DB 21; Length 1157;
 Best Local Similarity 62.0%; Pred. No. 1.4e-64;
 Matches 415; Conservative 0; Mismatches 251; Indels 3; Gaps 1;
 QY 1 ATGGTCTCTGGAGATTGAGATTAAGAGGATTGAAATACAAATCTAGAACCGACAGTCACT 60
 DB 179 ATGGGAAGAGAGAAATGAGATTAGAGGATTGAAATACAAATCTAGTCAAGTCACT 238
 QY 61 TTCTCAAGCGCCGAATGTTTATTAAGAAGCGTATGAATATCAGTTCTTTGTGAT 120
 DB 239 TTCTGTAAGAGAAATGGGTTGCTTAAAGAGCTTATGAATTTCTGTTCTTTGTGAT 298
 QY 121 CGAGAGTGGCCCTCATGCTCTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACAC 180
 DB 299 GCTGAAGTTGCTCTCATGCTTTCTCAAGCGGTGGCGCTCTATGAATTTGTAAACAC 358
 QY 181 AGCGTGAAGAGGAGATTGAGAGGTACAAGAGACTTGGCTTGACAAACACCGAGGG 240
 DB 359 AGTGTGAAGGCAACAATGATAGATATAAGAAAGCATCTCAGATTCTCCACACTGGA 418
 QY 241 CGATATCAGAGTCCCAATTTCTCAGTATTGGCAACAGGAGGCTGTGAATCTCAGAACAG 300
 DB 419 TCTACTTCTGAAGCTAACACTCAGTTTATCAACAGAGAGCTGCCAACTCCGAGTTGAG 478
 QY 301 ATTGACATTTTGAAGAGATGAGCAATTTGATGGGTGACCGGCTTACAGTTTGAAC 360
 DB 479 ATTGTAACCTTACAGAACTCAACAGGAACATGCTAGCGAGTCTCTAAGTTCTCTGACT 538
 QY 361 ATTAAGGAACCTCAAGCAACTTGAGGTTGACTTTGAAAGGAATCAGCGGAGTGGATCC 420
 DB 539 GCAAAAGATCTGAAGGCTGGAGACCAACTTGAAGGAATTTAGTAGAATTAGGTCC 598
 QY 421 AAAAAAGAGAGATGTTGCTTTGAAGAGATCGACATCATGAGAGAGAGGGAACATACCTT 480
 DB 599 AAAAAAGATGAACTCTCTGTTGCTTAAAGAGCTTATGAATTTCTGTTCTTTGTGAT 658
 QY 481 ATCCAGGAGATGAGATTTCTCGCAGCAGATAGCGGAGTGTGAGATAGCCACACAG 540
 DB 659 CACAACACATCAGATGCTTGGGCAAGATAGCTGAGAGTGAAGAAATGTGAACATG 718
 QY 541 AACATGTTATCAGCTCCGGAATATGATGCACTGCGCGCATTCGACTCTCGAAATTTCTTA 600
 DB 719 ---ATGGAGGAGAAATTTGAGTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 775
 QY 601 CATGCAATCTAATCGATGGCGCCATCATCTATGACATCAGGAACAAACAGCTTTCAG 660
 DB 776 CAAGTGAAGCGCTTACAGCAATATCATCAATATCCAGCCCAAGACACATGGCTCTTCAA 835

QY 661 CTTGGCTGA 669
 Db 836 TTAGTATAA 844

RESULT 12
 AAF85402
 ID AAF85402 standard; cDNA; 714 BP.
 XX AC AAF85402;
 XX AC AAF85402;
 DT 23-JUL-2001 (first entry)
 XX Nucleotide sequence of the floral homeotic protein PTAG-2.
 XX
 KW Floral homeotic gene; PTD; PTLF; PTAG-1; PTAG-2; floral tissue; LEAFY;
 KW LFY; FLORICAULA; FLO; DEFICIENS; DEF; AGAMOUS; AG; transgenic plant;
 KW fertility; sterility; ss.
 XX
 OS Populus balsamifera.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 1..714
 FT /*tag= a
 FT /product= "PTAG-2"
 XX
 XX CA2319853-A1.
 XX
 XX 01-APR-2001.
 XX
 XX 02-OCT-2000; 2000CA-2319853.
 XX
 XX 01-OCT-1999; 99US-0410464.
 XX
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 PI Rottman WH, Strauss SH, Brunner AM, Sheppard LA;
 XX
 DR WPI; 2001-336098/36.
 DR P-PSDB; AAB68438.
 XX
 XX Novel isolated polynucleotide derived from Populus species, useful for
 PT producing transgenic plants having modified fertility characteristic,
 PT particularly sterility -
 XX
 PS Claim-25; Page 61-62; 69pp; English.
 XX
 CC The present sequence encodes a floral homeotic protein, designated
 CC PTAG-2, which is derived from Populus balsamifera subsp. trichocarpa.
 CC The specification also describes PTD, PTLF, and PTAG-1 proteins. The
 CC floral homeotic proteins are expressed in floral tissues. PTLF is a
 CC homologue of LEAFY (LFY) and FLORICAULA (FLO), and is expressed in
 CC immature inflorescences on which floral primordia are developing. PTD
 CC is a homologue of DEFICIENS (DEF), and is strongly expressed in stamen
 CC primordia from the onset of organogenesis. PTAG-1 and PTAG-2 are
 CC homologues of AGAMOUS (AG). The floral homeotic proteins and
 CC polynucleotides are useful for producing transgenic plants having
 CC modified fertility characteristics, particularly sterility.
 XX
 XX Sequence 714 BP; 223 A; 146 C; 172 G; 173 T; 0 other;
 SQ

Query Match 27.3%; Score 248.2; DB 22; Length 714;
 Best Local Similarity 67.0%; Pred. No. 9.9e-64;
 Matches 352; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 2 TGGTCTGGGAGATTGAGATAAAGAGGATTGAAATATATAGAACCGACAGTCACTT 61
 Db 47 TGGGGAGGGGAAAGGTGAGATCAAGCGGATCGGAGAACCCACCAATCGCCAAAGTCACTT 106
 QY 62 TCTCAAGCGCCGAAATGGTTATTAAAGAGGGGTATGAATTATCAGTCTTTTGTGATG 121
 Db 107 TCTGCAAAAGCGGAAATGGTTTGTCTCAGAAAGCCTATGAATTATCTGTTTGGATG 166

QY 122 CAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTATGAATTTCCCAACCACA 181
 Db 167 CTGAGTTTGCACTCATCGTCTTCTCCAGCCCTTGACGCCCTTATGAGTACTTACACATA 226
 QY 182 CGGTGAAGAGACGATTGAGAGGTACAGAAGACTTGGTTTGACAAACACCGAGGGG 241
 Db 227 GTGTCAAATCTACAATTTGAAAGGTACAAAAGGCGATGTGCAGATTCTTTCCAAACACCGGT 286
 QY 242 CGATATCAGAGTCCAAATTTCTAGTATTGGCAACAGGAGGCTGGTAAACTCAGACAAACAGA 301
 Db 287 CAGTTTCTGAAGCCCAATGCTCAGTTCTATCAGCAGAGAGCTGCCAAGTGGCTGGCGCAA 346
 QY 302 TTGACATTTGCAAAATAGGCAATTTGATGGGTGACGGCTTACAGCTTTGAACA 361
 Db 347 TTGGTAAATTTGAGAAATTTCAACAGGAATATGCTGGGTGAATCACTTAGTCAATTGAGTG 406
 QY 362 TTAGCACTCAAGCACTTGAAGTTGAGTTGAAAGGAAATCAGCCGAGTGGATCCA 421
 Db 407 TGAAGAACTTAAAGAGCTTGGAGATAAAACTTGAGAAAGGAATTTGTTGAATTTCTTGA 466
 QY 422 AAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATGCGAGAGAGGGAACACATACCTTA 481
 Db 467 AAAAGAAATGAGCTGTTGTTGCTGAAATTTAGTATATGCAAGAGAGGAGATTGACTTGC 526
 QY 482 TCCAGGAGAAATGAGATTTCTTCCAGCAGCAGATGAGCGAGTGTGAGA 526
 Db 527 ACAACAATAACCAAGCTTCTCCGAGCAAAAGATTGCAAGAAATGAAA 571

RESULT 13
 AAZ57949
 ID AAZ57949 standard; cDNA; 1115 BP.
 XX AC AAZ57949;
 XX 11-APR-2000 (first entry)
 DE Poplar floral homeotic gene PTAG-2 cDNA.
 XX
 KW Poplar; PTAG-2; floral homeotic gene; transgenic plant; sterility;
 KW fertility; ss.
 XX
 OS Populus balsamifera subsp. trichocarpa.
 XX
 FH Key Location/Qualifiers
 FT CDS 99..815
 FT /*tag= a
 FT /note= "the coding region is also specifically
 claimed in Claim 1"
 XX
 XX CA2227940-A1.
 XX 06-OCT-1999.
 XX 07-APR-1998; 98CA-2227940.
 XX 06-APR-1998; 98US-0080851.
 XX (UYOR-) UNIV OREGON HEALTH SCI.
 XX
 XX Rottmann WH, Brunner AM, Sheppard LA, Strauss SH;
 XX WPI; 2000-106562/10.
 DR P-PSDB; AAY58657.
 XX
 PT Nucleic acid from Populus trichocarpa genes, useful for producing
 PT transgenic plants, particularly trees, with modified fertility
 PT characteristics such as sterility -
 XX
 PS Claim 1; Page 74-75; 92pp; English.
 XX
 CC The present sequence is that of cDNA encoding Populus balsamifera

CC subsp. trichocarpa PTAG-2 protein (see AY58657). The PTAG-2 gene (see
 CC AAZ57948) is 1 of 4 newly identified floral homeotic genes from this
 CC poplar species. PTAG-2 is a homologue of AGAMOUS and is expressed
 CC in floral tissue. The invention provides nucleic acid sequences of
 CC these 4 Populus genes, the corresponding cDNA sequences (see
 CC AAZ57942-49) and deduced amino acid sequences (see AY58454-57). It also
 CC provides methods of using the genes and cDNA to produce genetically
 CC engineered Populus species and other trees having modified fertility
 CC characteristics, including sterility. Genetic constructs useful in
 CC producing these genetically engineered trees include antisense
 CC versions of PTAG-2, dominant negative mutants, and constructs useful
 CC for sense suppression. Promoter sequences may be used to obtain
 CC floral specific expression of genes such as cytotoxins that may be
 CC used in genetic ablation strategies to produce trees having modified
 CC fertility characteristics, including sterility. Sterile trees
 CC allow increased wood yield and a reduction in the production of
 CC allergens such as pollen.

XX Sequence 1115 BP; 342 A; 240 C; 234 G; 299 T; 0 other;

Query Match 27.3%; Score 248.2; DB 21; Length 1115;
 Best Local Similarity 67.0%; Pred. No. 1.2e-63;
 Matches 352; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 2 TGGTCTGGTGGAGATTGAGATTAAGAGGATTGAAATACTACGACCGACAGGTCACATT 61
 DB 145 TGGGGAGGGGAGAGGTGGAGATCAAGCGATCGAGAACACCAATCGCCAGTCACATT 204
 QY 62 TCTGCAAGCCGCCGAATGGTTTATTAAAGAGCGGTATGAATATCAGTTCTTTGTGATG 121
 DB 205 TCTGCAAAAGGGCGGAATGGTTTCTCAAGAAAGCCATGAATATCTGTTCTTGGATG 264
 QY 122 CAGAAGTGGCCCTCATCTCTCTCCAGCAGAGGGAGATTATGAAATTTGCCAACCA 181
 DB 265 CTGAGGTGACATCATCTCTCTCTCCAGCGGTGGAGCCCTTATGAGTACTTAACAATA 324
 QY 182 GCGTGAAGGAGCGATTGAGAGGTACAGAAGACTTCGGTTGACACACACCGAGGGG 241
 DB 325 GTGTCAAAATCTACAATTGAAAGGTACAAAAGGCATGTGCAGATTCTTCCACAACGGGT 384
 QY 242 CGATATCAGAGTCCAAATCTCAGTATTGGCAACAGAGCGTGTGTAACCTCAGACACAGA 301
 DB 385 CAGTTCTGAAGCAATGCTCAGTTCTATCAGCAAGAGCTGCCAGCTCGCTCGCAA 444
 QY 302 TTGACATTTTGCAAAATGCAAAATAGGATTTGAGTGGTGAACGGGCTTACAGCTTTGAACA 361
 DB 445 TTGTAATTTGCAGAAATTCACACAGGAATATGCTGGGTGAATCATTGATGCTAGTGGATG 504
 QY 362 TTAAGGAACCTACGCAACTTGGGTTGAGTTCGACTTCAAAAAGCAATCAGCCGAGTCCGATCCA 421
 DB 505 TGAAGGAACCTTAAGAGCGTTGGAGATAAACTTGAAGAAAGGATTTGTAATTCGTTGCA 564
 QY 422 AAAAGAACGAGATGTTGCTTTGAAGAGATCGACATCATGCAAGAAAGGAAACACATCTTA 481
 DB 565 AAAAGATGAGCTGTTGCTTTGCTGAAATTCAGTATATGCAAGAGAGGAGATTGACTTGC 624
 QY 482 TCCAGGAGATGAGATTTCTCGCAGCAGAGATACCGAGTGTGAGA 526
 DB 625 ACAACAATACCAAGCTTTCTCCGAGCAAGAGATTGCAGAGATGAAA 669

RESULT 14

ID AAF85401 standard; cDNA; 1159 BP.

XX AAF85401;

AC AAF85401;

XX 23-JUL-2001 (first entry)

DE Nucleotide sequence of the floral homeotic protein PTAG-2.

XX Floral homeotic gene; PTD; PTLF; PTAG-1; PTAG-2; floral tissue; LEAFY;

KW LFY; FLORICAULA; FLO; DEFICIENS; DEF; AGAMOUS; AG; transgenic plant;

KW fertility; sterility; ss.

XX Populus balsamifera.

XX Key Location/Qualifiers

CDs 99..815

ET /*tag= a

FT /product= "PTAG-2"

XX CA2319853-A1.

XX 01-APR-2001.

XX 02-OCT-2000; 2000CA-2319853.

XX 01-OCT-1999; 99US-0410464.

XX (UYOR-) UNIV OREGON HEALTH SCI.

XX Rottman WH, Strauss SH, Brunner AM, Sheppard LA;

XX WPI; 2001-336098/36.

XX P-PSDB; AAB68438.

XX Novel isolated polynucleotide derived from Populus species, useful for
 XX producing transgenic plants having modified fertility characteristic,
 XX particularly sterility -

XX Claim 25; Page 60-61; 69pp; English.

XX The present sequence encodes a floral homeotic protein, designated
 XX PTAG-2, which is derived from Populus balsamifera subsp. trichocarpa.
 XX The specification also describes PTD, PTLF, and PTAG-1 proteins. The
 XX floral homeotic proteins are expressed in floral tissues. PTLF is a
 XX homologue of LEAFY (LFY) and FLORICAULA (FLO), and is expressed in
 XX immature inflorescences on which floral primordia are developing. PTD
 XX is a homologue of DEFICIENS (DEF), and is strongly expressed in stamen
 XX primordia from the onset of organogenesis. PTAG-1 and PTAG-2 are
 XX homologues of AGAMOUS (AG). The floral homeotic proteins and
 XX polynucleotides are useful for producing transgenic plants having
 XX modified fertility characteristics, particularly sterility.

XX Sequence 1159 BP; 386 A; 240 C; 234 G; 299 T; 0 other;

Query Match 27.3%; Score 248.2; DB 22; Length 1159;
 Best Local Similarity 67.0%; Pred. No. 1.2e-63;
 Matches 352; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 2 TGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAAATCTACGAAACCGACAGTCACATT 61
 DB 145 TGGGGAGGGGAAAGGTGGAGATCAAGCGATCGAGAACACCAATCGCCAGTCACATT 204
 QY 62 TCTGCAAGCGCCGAAATGGTTTATTAAAGAGCGGTATGAATATCAGTTCTTTGTGATG 121
 DB 205 TCTGCAAAAGGGCGGAATGGTTTGTCTCAAGAAAGCCTATGAATATCTGTTCTTTGCGATG 264
 QY 122 CAGAAGTGGCCCTCATCTCTCTCCAGCAGAGGGAGATTGAAATTTGCCAACCA 181
 DB 265 CTGAGGTGCACTCATCTCTCTCCAGCGGTGAGCCCTTATGAGTACTTAACAATA 324
 QY 182 GCGTGAAGGAGCGATTGAGAGGTACAGAAGACTTCGGTTGACAAACACCGAGGGG 241
 DB 325 GTGTCAAAATCTACAATTGAAAGGTACAAAAGGCATGTGCAGATTCTTCCACAACGGGT 384
 QY 242 CGATATCAGAGTCCAAATCTCAGTATTGGCAACAGAGCGTGTGTAACCTCAGACACAGA 301
 DB 385 CAGTTCTGAAGCCAAATGCTCAGTTCTATCAGCAAGAGCTGCCAGCTCGCTCGCAA 444
 QY 302 TTGACATTTTGCAAAATGCAAAATAGGCAATTTGAGTGGTGAACGGGCTTACAGCTTTGAACA 361
 DB 445 TTGTAATTTGCAGAAATTCACACAGGAATATGCTGGGTGAATCATTGATGCTAGTGGATG 504
 QY 362 TTAAGGAACCTACGCAACTTGGGTTGAGTTCGACTTCAAAAAGCAATCAGCCGAGTCCGATCCA 421

Db 505 TGAAGGAACCTTAAGAGCTTGAGATAAACTTGAGAAAGGAATTGTAGAAATTCGTTTGA 564
Qy 422 AAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATGAGAGAAAGGGAACACATCTTA 481
Db 565 AAAAGAAATGAGCTGTTGTTGCTTGAATGAGTATATGAGAGAGGGAGATTGACTTGC 624
Qy 482 TCAGAGAGATGAGATCTTCGAGAGAGAGATAGCCGAGTGTCA 526
Db 625 ACAACAATAACAGCTTCTCCGAGCAAGATTGAGAGAAATGAAA 669

RESULT 15

ABK88491
ID ABK88491 standard; cdna; 1159 BP.

XX AC ABK88491;

XX DT 07-OCT-2002 (first entry)

XX DE Poplar floral homeotic protein, PTAG-2, cdna.

XX KW Poplar; ss; gene; plant; DEFICIENS; transgenic; promoter;
KW protein transduction domain; floral homeotic gene;
KW floral-specific expression; cytotoxin; fertility; sterility;
KW PTLF; PTAG-1; PTAG-2.

XX OS Populus balsamifera subsp. trichocarpa.

XX FH Key Location/Qualifiers
FT CDS 99..815
FT /*tag= a
FT /*product= "PTAG-2"

XX PN US6395892-B1.

XX PD 28-MAY-2002.

XX XX 01-OCT-1999; 99US-0410464.

XX XX 06-APR-1998; 98US-080851P.

XX XX 06-APR-1999; 99US-0287700.

XX XX (UYOR-) UNIV OREGON HEALTH SCI.

XX XX Strauss SH, Rottmann W, Brunner A, Sheppard L;

XX XX WPI; 2002-572853/61.

XX XX P-PSDB; ABG30866.

XX PT New protein transduction domain promoter nucleic acid molecule useful
XX for producing transgenic plants having modified fertility
XX characteristics, particularly sterility -
XX Disclosure; Column 77-80; 46pp; English.

XX CC The invention relates to an isolated nucleic acid molecule especially a
XX protein transduction domain (PTD) promoter: (i) that hybridises under
XX wash conditions of 0.2 x SSC (saline sodium citrate), 0.1 %SDS (sodium
XX dodecyl sulphate) at 65 plusC to nucleotides or (ii) comprising 35
XX consecutive nucleotides of the PTD gene. PTD is a floral homeotic
XX gene and is the homologue of DEFICIENS. Also includes are a recombinant
XX nucleic acid comprising the PTD promoter, a cell transformed with
XX the recombinant nucleic acid and a transgenic plant comprising the
XX transformed cell. The PTD promoter is useful to obtain floral-specific
XX expression of genes such as cytotoxins, that are employed in genetic
XX ablation strategies to produce trees having modified fertility
XX characteristics, including sterility. Genetic constructs comprising
XX antisense versions or dominant negative mutants of PTD are useful in
XX producing genetically engineered Poplars and other trees, and for sense
XX suppression. Also disclosed are 3 other homeotic genes PTLF, PTAG-1
XX and PTAG-2 (none are defined). The present sequence is the PTAG-2 cdna.

SQ Sequence 1159 BP; 386 A; 240 C; 234 G; 299 T; 0 other;
Query Match 27.3%; Score 248.2; DB 24; Length 1159;
Best Local Similarity 67.0%; Pred. No. 1.2e-63;
Matches 352; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

Qy 2 TGGGTGCTGGCAAGATTGAGATAAAGAGGATTGAAATATCTACCAACGACAGTCACTT 61
Db 145 TGGGAGGGGAAAGGTGGAGATCAAGCGGATCGAGAACACCAACATCGCCAAATGCTT 204
Qy 62 TCTGCAAGCGCGGAAATGGTTTATTAAAGAGGCGGTATGAATTTATCATGTTTGTGATG 121
Db 205 TCTGCAAAAGGCGGAATGGTTTGTCTCAAGAAAGCCTATGAATTTATCTGTTTTCGATG 264
Qy 122 CAGAAAGTGGCCCTCATGCTCTCTCCAGCAGAGAGGAGACTTTATGAATTTGCCAACACCA 181
Db 265 CTGAGGTTGCACTCATGCTCTCTCCAGCCTGACCGCTTTATGAGTACTCTTAAACATA 324
Qy 182 CGGTGAAGAGGACGATTGAGAGGTACAAGAGACTTGCCTTGACAAACACCGAGGGG 241
Db 325 GTGTCNAATCTACAATTGAAAGGTACAAAAGGCGATGTGCAGATTCTTCCAACAACGGGT 384
Qy 242 CGATATCAGATCCAAATCTCAGTATTGGGCAACAGGAGGCTGGTAAACTCAGACAAACAGA 301
Db 385 CAGTTTCTGAAAGCCAATGCTCAGTTCTATCAGCAAGAAAGCTGCCAAGCTGCGTCGCAAA 444
Qy 302 TTGACATTTTGC AAAAATGCAAAATAGGCAATTTGATGGGTGACGGGCTTACAGCTTTGAACA 361
Db 445 TTGTAATTTGCAGAAATTCAAAACAGGAATATGCTGGGTGATCACTTAGTGCAATTGAGTG 504
Qy 362 TTAAGGAACTCAAGCAACTTTCAGGTTTCGACTTTGAAAAAGGAATCAGCGGAGTGGATCCA 421
Db 505 TGAAGGAACCTTAAGAGGCTTGGAGATAAAACCTTGAGAAAGGAATTTGGTAGAATTCGTTGA 564
Qy 422 AAAGAACGAGATGTTGCTTGAAGAGATCGACATCATGCAAGAAAGGGAACACATACATA 481
Db 565 AAAGAAATGAGCTGTTGTTTCTGAAATTTGAGTATATGCAAGAGGGAGATTGACTTGC 624
Qy 482 TCCAGGAGATGAGATTCTTCGAGCAAGATAGCCGAGTGTCA 526
Db 625 ACAACAATAACAGCTTCTCCGAGCAAGATTGAGAGAAATGAAA 669

Search completed: January 21, 2004, 17:31:51
Job time : 238.397 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 15:50:10 ; Search time 3690.68 Seconds
(without alignments)
11217.584 Million cell updates/sec

Title: US-09-936-869-5
Perfect score: 1012
Sequence: 1 atcgatttaagtcattggc.....ttctatggttaaaaaaaaaa 1012

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2888711 seqs, 20454813366 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba:
2: gb_htg:
3: gb_in:
4: gb_in:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_un:
14: gb_vi:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_or:
21: em_ov:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_un:
29: em_vi:
30: em_htg_hum:
31: em_htg_inv:
32: em_htg_other:
33: em_htg_mus:
34: em_htg_pln:
35: em_htg_rod:
36: em_htg_mam:
37: em_htg_vrt:
38: em_sy:
39: em_htgo_hum:
40: em_htgo_mus:
41: em_htgo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1012	100.0	1012	8	ATHRNS2X	M98336 Arabidopsis
2	924	92.0	1208	8	AB032256	AB032256 Nicotiana
3	294.4	29.1	1045	8	AM0315591	AJ315591 Anirrhin
4	261.2	25.8	92624	8	AC003000	AC003000 Arabidops
5	249	24.6	9423	8	ATH275979	AJ275979 Arabidops
6	182.8	18.1	981	8	AF139660	AF139660 Calystegi
7	102.2	10.1	4210	8	AB112027	AB112027 Nicotiana
8	82.6	8.2	105384	8	AF008430	AF008430 Lotus jap
9	78.4	7.7	4937	8	AWA489249	AU489249 Antirrhin
10	55.2	5.5	56284	8	AC007195	AC007195 Arabidops
11	54.6	5.4	142955	8	AF003349	AF003349 Oryza sat
12	54.6	5.4	145265	8	AF003418	AF003418 Oryza sat
13	53.4	5.1	78756	8	PFMAL6P1_13	Continuation (14 o
14	53.4	5.1	110000	2	PFMAL6P1_12	Continuation (13 o
15	50.8	5.0	103098	8	ATT5C2	AL138664 Arabidops
16	50.4	5.0	1030	8	AY086747	AY086747 Arabidops
17	50.4	5.0	2000	6	AX655393	AX655393 Sequence
18	49.2	4.9	188177	2	BX119918	BX119918 Danio rer
19	47.2	4.7	158475	9	AC097463	AC097463 Homo sapi
20	46.8	4.6	693	6	AX412377	AX412377 Sequence
21	46.8	4.6	693	6	AX412567	AX412567 Sequence
22	46.8	4.6	693	6	AX507176	AX507176 Sequence
23	46.8	4.6	763	8	AY114710	AY114710 Arabidops
24	46.8	4.6	881	8	AY072413	AY072413 Arabidops
25	46.8	4.6	925	8	ATU05206	U05206 Arabidopsis
26	46.8	4.6	1141	6	AX083744	AX083744 Sequence
27	46.8	4.6	223954	2	AC106472	AC106472 Rattus no
28	46.8	4.6	234320	2	AC120569	AC120569 Rattus no
29	46.8	4.6	261007	2	AC133419	AC133419 Rattus no
30	46.2	4.6	234894	2	AC111286	AC111286 Rattus no
31	46.2	4.6	265793	2	AC115477	AC115477 Rattus no
32	45.6	4.5	273275	3	AE014828	AE014828 Plasmodiu
33	45.4	4.5	935	8	CAR012689	AE012689 Cicer ari
34	45.2	4.5	177363	9	AC073856	AC073856 Homo sapi
35	45.2	4.5	201359	2	AC128640	AC128640 Homo sapi
36	45	4.4	233746	2	AC120480	AC120480 Rattus no
37	45	4.4	270166	2	AC108263	AC108263 Rattus no
38	45	4.4	270972	2	AC132556	AC132556 Rattus no
39	44.8	4.4	331	8	AF504288	AF504288 Crataegus
40	44.6	4.4	506	8	AY260449	AY260449 Prunus ar
41	44.4	4.4	505	6	BD119695	BD119695 EST and e
42	44.2	4.4	125020	9	AF429315	AF429315 Homo sapi
43	44.2	4.4	192581	2	PFMAL13P1	AL049180 Plasmodiu
44	44	4.3	894	8	LERNALE	X79337 L.esculentu
45	43.6	4.3	21024	1	AE014269	AE014269 Streptoco

ALIGNMENTS

RESULT 1

ATHRNS2X

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

ATHRNS2X

Arabidopsis thaliana

M98336

M98336.1

ribonuclease.

Arabidopsis thaliana

Eukaryota; Viridiplantae;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Xosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi

1 (bases 1 to 1012)

Taylor,C.B., Baricola,P.A., delCardayre,S.B., Raines,R.T. and

Green,P.J.

1012 bp

mRNA

linear

PLN 30-OCT-1994

complete cds.

1012 bp

mRNA

linear

PLN 30-OCT-1994

complete cds.

TITLE RNS2: a senescence-associated RNase of Arabidopsis that diverged from the S-RNases before speciation
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90 (11), 5118-5122 (1993)
MEDLINE 93281708
PUBMED 8506358
COMMENT Original source text: Arabidopsis thaliana cDNA to mRNA.
FEATURES Location/Qualifiers
 1..1012
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 1..1012
 /gene="RNS2"
 16..795
 /gene="RNS2"
 /codon_start=1
 /product="ribonuclease"
 /protein_id="AA51406.1"
 /db_xref="GI:299210"
 /translation="WASRLICLLLVACIAGAPADVIELNRSQREPPYFALSLOWPST YKGRHCCSKNACCGSDATQFTIHLNPDYNDGWPSCCYRSDPKREKISTLMDG LKGYWPSLSCGSPSCNGKGSFWEHWEKHGTCSPVFHDEYNYFLTINLYLKGV TTVLYOAGYVANSSEKYPILGGIVTAIONAPHITPEVVKRDAIDEIRICFYKDFKPRD CVGSDLTRSKCPKYVSLPVLPTLDGEAMVLKMPETERAL"
BASE COUNT 287 a 197 c 217 g 311 t
ORIGIN
 Query Match 100.0%; Score 1012; DB 8; Length 1012;
 Best Local Similarity 100.0%; Pred. No. 6.8e-247; Indels 0; Gaps 0;
 Matches 1012; Conservative 0; Mismatches 0;
 1 ATCGAATTAAGTCAATGCGTCACGTTTATGCTTCTCTCCCTTCGTTGGTGTATGCC 60
 2 ATCGAATTAAGTCAATGCGTCACGTTTATGCTTCTCTCCCTTCGTTGGTGTATGCC 60
 3 GGAGCATTTGCCGGAGAGCTGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 120
 4 GGAGCATTTGCCGGAGAGCTGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 120
 5 GCTCTATCTCTTCAATGGCTGGAACTATTGCCGTGGAACCTGCCATTTGTGTCCAAA 180
 6 GCTCTATCTCTTCAATGGCTGGAACTATTGCCGTGGAACCTGCCATTTGTGTCCAAA 180
 7 AACGTTGCTGAGAGGCTCCGATGCTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 240
 8 AACGTTGCTGAGAGGCTCCGATGCTCAATCAATCAATCAATCAATCAATCAATCAATCAAT 240
 9 GACTATACCATGTTTGTGGCTTCATGTTGTTATCGATCTGACTTTAAAGAGAAGAG 300
 10 GACTATACCATGTTTGTGGCTTCATGTTGTTATCGATCTGACTTTAAAGAGAAGAG 300
 11 ATTTCACCTTCATGATGGTCTTGGAGAGTACTGGCTAGTCTCAGTTGTTGTTCTCCA 360
 12 ATTTCACCTTCATGATGGTCTTGGAGAGTACTGGCTAGTCTCAGTTGTTGTTCTCCA 360
 13 TCATCATGCAATGTTGGAAAGGTCATTTTGGGCCACGAGTGGGAGAACATGGGACT 420
 14 TCATCATGCAATGTTGGAAAGGTCATTTTGGGCCACGAGTGGGAGAACATGGGACT 420
 15 TGTCTCTCTCTGTTTTCATGATGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 16 TGTCTCTCTCTGTTTTCATGATGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 17 TTGAAGCATTAATGTCACGATGCTCTTTATCAAGCTGGCTATGTTGTTCCAAAGTGAA 540
 18 TTGAAGCATTAATGTCACGATGCTCTTTATCAAGCTGGCTATGTTGTTCCAAAGTGAA 540
 19 AAGTATCTCTAGGAGGTATGTAACAGCCATTCAGATGATTTATATATATATATATATATATAT 600
 20 AAGTATCTCTAGGAGGTATGTAACAGCCATTCAGATGATTTATATATATATATATATATATAT 600
 21 GTGGTTGCAAAAGAGATGCAATCGATGAATACGATATGCTTCTATTAAGATTTTAAG 660

RESULT 2
AB032256 LOCUS 1208 bp mRNA linear PLN 17-SEP-1999
DEFINITION Nicotiana glutinosa NGR2 mRNA for RNase NGR2, complete cds.
ACCESSION AB032256
VERSION AB032256.1 GI:5902453
KEYWORDS RNase NGR2,
SOURCE Nicotiana glutinosa
ORGANISM Nicotiana glutinosa
REFERENCE 1 (sites)
AUTHORS Kariu, T., Kurata, N., Yamasaki, N. and Kimura, M.
TITLE Molecular cloning and characterization of cDNAs encoding ribonucleases in Nicotiana glutinosa leaves that are induced by wounding and TMV-infection
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1208)
AUTHORS Kimura, M.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-1999) Makoto Kimura, Kyushu University, Faculty of Agriculture, Hakozaki 6-10-1, Higashi-Ku, Fukuoka, Fukuoka 812-8581, Japan (E-mail:mkimura@agr.kyushu-u.ac.jp, Tel:81-92-642-2854 Fax:81-92-642-2854)
FEATURES Location/Qualifiers
 1..1208
 /organism="Nicotiana glutinosa"
 /mol_type="mRNA"
 /db_xref="taxon:35889"
 1..1208
 /gene="NGR2"
 1..1208
 /gene="NGR2"
 72..905
 /gene="NGR2"
 /codon_start=1
 /product="RNase NGR2"
 /protein_id="BAA84468.1"
 /db_xref="GI:5902454"
 /translation="MTPLPVHATVNLRLMLPLIGACILGVNGGNDGVEGLLRGRG GAGGYQREDFEKLISLQWPGTCTRETRCCSNACCSNSPPVFTIHLWTEYNDG TWPACCKRAPDEREISLLEPKYKWPISLSCSPSCRRKGRKGFPAHWEKHGTCAV PVHDEYEFELTINLYFYKNFVPEVSEKYPILGGIISINAPHATPEL TCSGDALDEELRICFYKNFEPDRDCARDTSALSRGSCFPQVSLPAHGSWEISIRISNTAA S"

```

BASE COUNT      359 a      232 c      270 g      347 t
ORIGIN
Query Match      32.0%; Score 324; DB 8; Length 1208;
Best Local Similarity 67.5%; Pred. No. 1.2e-71;
Matches 456; Conservative 0; Mismatches 220; Indels 0; Gaps 0;
QY 73 GGAGAGTCATCGAATCAATCAATCATCTCAGAGGAGTTGCGATTATTCGCTCTATCTCT 132
Db 195 GGAAGCGGGCGCGCAGAGGAGATATCAGCGAGAGTTGCGATTATTCGCTCTATCTCTC 254
QY 133 CAATGCGCTTGAACCTATTGCGGTGAACCTGCCCATTTGCTCCAAAACGCTTGCTGC 192
Db 255 CAATGCGCGAGGACCTATTGCGGAGAGACCCGCGCTTGTGTTCTTCAGCGCTGTTC 314
QY 193 AGAGGCTCGATGCTCCAACTCAATCAATCAATCAATGATGGTTATGGCTCTGATCAACGAT 252
Db 315 AGCGGCTCGAATCGCCACCTGATTTCAATTCATGACATTTGGACAGAGTACAAATGAT 374
QY 253 GGTTCGTGGCTTCATGTTGTTATCGATCTGACTTTTAAAGAGAGGAGATTTCAACGTTG 312
Db 375 GGAACCTTGGCTGCGCTGCTGTTCCGGTAAAGCATTTGATGAGAGAGATTTCAACATG 434
QY 313 ATGATGCTCTTGAAGATCTAGGCTAGTCTCAGTTGTTGTTCTTCCATCATCAATCAAT 372
Db 435 CTCGAGCCCTTGAGGAAAGTATTGGCTCTTTTAACTTGGTGTCCCGCAGATCTTGGCAT 494
QY 373 GGTGGCAAAAGGTCATTTTGGGCGCCAGAGTGGGAGAACATGGGACTTGTCTTCTCCT 432
Db 495 CACAGAAAGAGACCATCTGGGCATATGATGGGAAAACATGGGACATGTGCTTATCCA 554
QY 433 GTTTTCATGATGAGTAAATTAATTCCTTCCACCATTAATCTCTACTTTGAAGCATPAAT 492
Db 555 GTTGTCCATGATGAATATGATGCTTTTGTGATACATGAATATTTACTTCAAGTATAAT 614
QY 493 GTCACGATGTCCTTTATCAAGCTGCTATGTTGCTTCCACAGTGAAGTAAAGTATCCTCTA 552
Db 615 GTTACAGAAGTCTGTTGAAGCTGATATGATCAATCAGATTCGAAAAGTATCCATTA 674
QY 553 GGAGTATCGTAACAGCCATTCAGAAATGATTTCAATATCAACCTGAGTGGTTGCAAA 612
Db 675 GGAGGATCATTTTCATCAATTTGAAACGCTTTTCATGCAACCCAGAGTTTAAATGTCTA 734
QY 613 AGAGATGCAATCGATGAATAGTATATGCTTCTATTAAGATTTTAAAGCCAGGAGCTGT 672
Db 735 GCGATGCTCTGAGGAACTTCTGATATGCTTCTATAGAAATTTTGAGCTCTGATGTGT 794
QY 673 GTTGGTTCAAGATTTGACATCTAGAAAGTCATGCCCAAGTACGTAAGTTTGGCGAA 732
Db 795 GCACGATPACTAGCGCTTATCAAGAGGTCATGCTCCTCAATATGTCAGCTTGCCAGCC 854
QY 733 TACACGCCATTAAGATG 748
Db 855 CATGGATCATGGATG 870

RESULT 3
AM0315591      1045 bp      mRNA      linear      PLN 10-AUG-2001
LOCUS      Antirrhinum mollissimum mRNA for S-like RNase 28 (s128 gene).
DEFINITION      AJ315591
ACCESSION      AJ315591
VERSION      AJ315591.1 GI:15149818
KEYWORDS      Ribonuclease; RNase.
SOURCE      Antirrhinum mollissimum
ORGANISM      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
      Asteridae; lamiids; Lamiales; Antirrhinaceae; Antirrhineae;
      Antirrhinum.
REFERENCE      1
AUTHORS      Liang, L. and Xue, Y.
TITLE      Molecular characterization of two S RNase-like genes from
      Antirrhinum

```

JOURNAL REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 1045)
Xue, Y.
Direct Submission
Submitted (11-JUN-2001) Xue Y., Plant Genetics and Developmental
Biology, Institute of Developmental Biology, CAS, 3 Nanyitiao,
Zhongguancun, Beijing 100080, CHINA

FEATURES
source

1..1045
/organism="Antirrhinum mollissimum"
/mol_type="mRNA"
/cuiVar="552"
/db_xref="taxon:162390"
/clones="PSL28"
/tissue_type="pistil"
/dev_stage="open flower"
33..863
/function="ribonuclease"
/codon_start=1
/product="S-like RNase 28"
/protein_id="CAC50874.1"
/db_xref="GI:15149819"
/translation="MAVLTARPLNPAIQACFCVILWIGLLCVNNGINGSDLGEKLG
ANQDFDFHIALOWPFTGRTTCHCCPTGCGCRGNAPAEFTIHGILWPDYNDGSPWS
CCTCKTPEKEISTLLGLNKPWSLSCGSPSNCHGKGLFWEHWEKHGTCTGSSVPS
AEYNFVTLKVPFKYNTVRLRAGVVASNSEKYPGLGIVTAIQAFHATPELKSG
DAVELYLCFKNFEPRDCATKSNKSCPRVSLPYSYLKMANDGNEVSEALSDSI
"

CDS

BASE COUNT 308 a 204 c 229 g 304 t
ORIGIN

Query Match 29.1%; Score 294.4; DB 8; Length 1045;
Best Local Similarity 67.4%; Pred. No. 4.2e-64;
Matches 434; Conservative 0; Mismatches 201; Indels 9; Gaps 1;

```

QY 100 CAGAGGAGTTGCGATTATTCGCTCTATCTCTCAATGGCTCGAACCTATTCGGTGA 159
Db 171 CAGAGAGACTTTGATTAATCTCCATCTCGCTCTCAATGGCCAGGACATTTCTCCCGAGA 230
QY 160 ACTGCGCATTTGCTCCAAAAACGCTTGTGTCAGAGGCTCCGATGTCCTCAACTCAATTC 219
Db 231 ACCGCGCATTTGTCGCCCAACCAATGGCTGTTGTCGAGGTTCAAACGACCGGAGAAATTT 290
QY 220 ACAATTCATGGTTATGCGCTGACTATAACGATGGTTGCGGCTTCATGTTGTTATCGA 279
Db 291 ACAATTCATGACTTTGGCTGATTAACAATGATGAAGTTGGGCTTCATGTTGCACTGGT 350
QY 280 TCTGACTTTAAAGAGAGAGAGATTTCAACGTTGATGGATGTTCTTGAGAGGACTCGGCT 339
Db 351 AAAAAGTTTGAGGAAAAGAGATTTCTACATTTGTTGACCTGAATAGTATTGCGCA 410
QY 340 AGTCTCAGTTGTTGTTCTCCATCATCATGCAATGGTGGGAAAGGTCATTTGGGCGCAC 399
Db 411 TCATTAAGCTGTGGTTCTCCGTCAAACTGCGATGGTGGGAAAGGATATTTTGGGAGCAT 470
QY 400 GAGTGGGAGAAACATGGGACTTTGTTCTTCCTGTTTTTTCATGATGAGTATTAATCTTC 459
Db 471 GAGTGGGAGAAACATGAACTTTGTTGCTTCATGTCACAGAGCTGAATACAAATTAATTC 530
QY 460 CTTACACACTTAATCTCTACTTGAAGCATAATGTCACGATGCTCTTTATCAAGCTGGC 519
Db 531 GTGACAGACTTAAAGTTTATTTCAATACAAATGTCACGGAAGTTCTGAGAGAGCTGGT 590
QY 520 TATGTTGCTTCCACAGTGAAGATTCCTCTAGGAGGTATCGTAACAGCCATTCAGAAAT 579
Db 591 TATGTAGCATCAAAATCTGAAAATATCTTTAGGAGGATCGTTACAGCAATACAAAT 650
QY 580 GCATTTCAATACCCCTGAGTGGTTTGCAAAAGAGATGCAATCGATCAATACATATATA 639
Db 651 GCTTTCATCGACTCCAGAACTTAATGCTCGGGTGACGCTGGGAGAACTTTACTTGT 710
QY 640 TGGTCTTATAAGATTTTAAAGCCAGGAGCTGTGTTGGTTTCAACAGATTGACATCTAGA 699

```

Db 711 TGCTTCTACAGAATTTTGAGCCTCGTGATTGTG-----CAACCAATCCACAAA 761

Qy 700 AAGTCATGCCCAAGTACGTAAGTTTGCCGGRATACAGCCATT 743

Db 762 AAATCGTGCTCCAGATATGTGAGCTTGCCAGATATTCCTCAT 805

RESULT 4

AC003000 Arabidopsis thaliana (thale cress)

LOCUS AC003000.3 GI:20196974

DEFINITION Arabidopsis thaliana chromosome 2 clone T517 map C1C10A06, complete sequence.

ACCESSION AC003000

VERSION AC003000.3

KEYWORDS HTG.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta, eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 92624)

AUTHORS Rounsley, S.D., Lin, X., Ketchum, K.A., Crosby, M.L., Brandon, R.C., Sykes, S.M., Kaul, S., Mason, T.M., Kerlavage, A.R., Adams, M.D., Somerville, C.R. and Venter, J.C.

JOURNAL Unpublished 2 (bases 1 to 92624)

REFERENCE 2 (bases 1 to 92624)

AUTHORS Lin, X.

TITLE Direct Submission

JOURNAL Submitted (09-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA

REFERENCE 3 (bases 1 to 92624)

AUTHORS Town, C.D. and Kaul, S.

TITLE Direct Submission

JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, cdtonet@igr.org

COMMENT On Apr 18, 2002 this sequence version replaced gi:6598383.

FEATURES

Location/Qualifiers

1..92624

organism="Arabidopsis thaliana"

molecule="genomic DNA"

cultivar="Columbia"

db_xref="taxon:3702"

chromosome="2"

map="C1C10A06"

clones="T517"

misc_feature 1..6205

note="overlap with BAC clone F17A14 (AC003674:32937..39141)."

complement(129..169)

repeat_region /rpt_family="CAAAA"n"

complement(1000..1024)

repeat_region /rpt_family="(A)n"

complement(1266..1291)

repeat_region /rpt_family="(CAAAA)n"

complement(2779..2799)

repeat_region /rpt_family="(A)n"

complement(2880..2934)

repeat_region /rpt_family="AT_rich"

3266..3290

repeat_region /rpt_family="AT_rich"

complement(4553..5912)

gene /gene="T517.1; similar to GP 2245012 gnl PID e327000 Z97341; supported by cDNA: gi_15293230_gb_AY051049.1"

complement(<4553..>5912)

mRNA /gene="At2g39710"

CDS complement(4584..5912)

gene /gene="At2g39710"

complement(4584..5912)

repeat_region /codon_start=1

product="expressed protein"

protein_id="AAB87120.2"

db_xref="GI:20196976"

/translation="MMDSSSSSLSSLSKFLRISVLLLIIFPLTFCKTSNTQTLLESL
KTQLPSSSDKLSFRHNVTITVLAAGDPQNIQSMVLDTGSELSELUHCKKSPNLGVS
FNPVSGTSYPCSSPICRTRTRDLIPASCDPKTHLCHVAISYADATIEGNLAHE
TFVGSVTRPTGLFGCNDLGLSSNSEDADKSTGLMNGRSLFVNQLGFSKFSYCS
GSDSGLLGLDASYSWLPQIOTPLVLOSTPLPFDYVAVTVOLEIRVGSKLISLP
KSVFPHDTGAGQTMDSGTQFTFLGMPVYALKNEFITQKTSVLRLEVDPDDFVQGT
MDLCYKGVSTIRNFSGPLVSLMFRGAEKMSVSGQKLLYVNGAGSGSKVEYCTEG
NSDLIGIEAFVIGHHQQNVNMFELAKSRVGFAGNVRCDLASRLGLRD"

complement(8942..8968)

repeat_region /rpt_family="AT_rich"

gene /gene="At2g39720"

complement(9879..11347)

note="synonym: T517.2; supported by full length cDNA: Ceres:253849"

mRNA complement(<9879..>11347)

CDS /gene="At2g39720"

complement(10039..11244)

gene /gene="At2g39720"

codon_start=1

product="expressed protein"

protein_id="AAB87121.1"

db_xref="GI:2642154"

/translation="MASGSVWCYSYCSRFVWVSDISCPDCGGLELIQBLDFTPSD
SFTTTTQHSPTFRFPFSSSTPSASMHADNSPTPIVTRIRSRSPNPVQLVIRG
SAAAPSSDVSEGLDSAFQMYDDGTDSGLRPLFPSPFLLGSGDFDLQISQIE
LNTNRNLRSCHEPPASAKIHALPLIIDPHTLLSDSCAVCKENFLKSSAREMP
CNHLYHPDCILPLAIRNSCPVCRHELPAEDLTDGTGAALTAVTATAEEDSDAAGLT
IWRIPGGGPAVGRIPIGGWRGDRMPVYVTEVCGRLGDRLPVAVGSRGRGDRGG
SSRRNGGFAGRIMRLFCGFSGSGSIAAAAASSGSRIRVTRTRTSFMSFSTASS
GSRNRNMLLA"

gene /gene="At2g39725"

12014..13082

note="synonym: T517.21; supported by full length cDNA: Ceres:5605"

mRNA join(12014..12116,12596..13082)

CDS /gene="At2g39725"

12603..12875

gene /gene="At2g39725"

codon_start=1

product="Expressed protein"

protein_id="AAM14851.1"

db_xref="GI:20196975"

/translation="MGASKLSGKQKVLRYGELRAARSPIEDRKRIEMIVSTEF
RNSKEVDKRNFGYIEYLLRLTKQLDQLKSPAIYVLSLSIKVVTSTK"

complement(13325..16081)

gene /gene="At2g39730"

note="synonym: T517.18; supported by full length cDNA: Ceres:7114"

mRNA complement(join(<13325..13599,13724..14005,14087..14559,14651..14736,14836..14941,15033..15349,15834..>16081))

CDS /gene="At2g39730"

complement(join(13484..13599,13724..14005,14087..14559,14651..14736,14836..14941,15033..15349,15834..>15878))

gene /gene="At2g39730"

codon_start=1

product="expressed protein"

protein_id="AAB87122.1"

db_xref="GI:2642155"

/translation="MAAAVSTVGAINRAPLSLNGSGSAGVSPASTFLGKVKVTVSRF
AOSNKKSGSFVKLVKEDKQTDGRWGLAYDSDQDITRGKGVDSVQAQPMGT
GTHAVLSYEVYQGLRQLNDLNMMDGYTAPAFMDLVVHTIKNLTLPNKKVPLI
LGIGWKGQKGFQCELVMAKMGINPINSAGELESNGAGPAKIRORYEAADLIK
KGMKCLFINDIDAGARGMGTTQYVANNQVNTLMIADNPTVQLPFGMYNKEENA
RPICTGNDSTLYPLIRDRGMEKFWAPTRDIRIGVCKIGITDKIKDIDIIVLV
DPPQGSIDIFGALRARIYDDVRKFEVSLGVEKIGKLVNSREGPPVPEQPEMTVEK
LMBYGNLMVEQENVRVOLAETLYLSQAALGANADAIGRTGYFGKGAQVNLVPPEG
CTDPVAENFDPTARSDDTGVYNF"

repeat_region /rpt_family="AT_rich"

17291..17350

repeat_region /rpt_family="AT_rich"

18075..18106

gene /rpt_family="AT_rich"

18783..21237

gene /gene="At2g39740"

/note="synonym: T517.4; predicted by gensean"
join(18783. .18921,19188. .19352,19496. .19674,19766. .19861,
19965. .20034,20131. .20261,20593. .>21237)
/gene="At2g39740"
CDS
join(18783. .18921,19188. .19352,19496. .19674,19766. .19861,
19965. .20034,20131. .20261,20593. .21237)
/gene="At2g39740"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAB87123.1"
/db_xref="GI:2642156"
/translation="MSRNPFLDPTLQELIQLVQIKPTRADRDTRITVIDQLDVLQSVEC
LRGATVQPFQSFVSNLFRWGLDIDVDLFSGSSILFTGKKOKOTLLGHLRLARASG
LWYKLOFVHARVPIILKVSGHORIISCDISIDNGLKSLRFLWSEIDGPRDLVL
LVKEMAKENINDSKTGTNSVLSLLVIFHQTCVPAILPLERIVPKSAVDLITGV
RKYAESIAQVTAARFARSKRAKSVRSSELLVFFAKVDFEOPVNAARSVS
RNLRIQAQVQITSRRLVSECNRSIIIGLTQHIQIESLYRTISLPQSHANGHWN
RNLHGQRQNOQMOQSQSYNTPNPFWPLTQSRPQONQNNRNLQOQPVQVQ
QWFWITQTQTKQSPYKSGNRPLKTSAGSSONQGHICKPSGHVNGVNSARPAYTNG
VNSARPSKIPSGOGQIMRPRHEQ"
/complement(21381. .24860)
/gene="At2g39750"
/note="synonym: T517.5; supported by cDNA:
GI_15450748.gb AY053416.1"
complement(join(<21381. .21726,21839. .21916,22124. .22246,
22328. .22585,22834. .22993,23212. .23322,23481. .23751,
23847. .23996,24089. .>24860))
/gene="At2g39750"
complement(join(21519. .21726,21839. .21916,22124. .22246,
22328. .22585,22834. .22993,23212. .23322,23481. .23751,
23847. .23996,24089. .24814))
/gene="At2g39750"
/codon_start=1
/product="expressed protein"
/protein_id="AAB87124.1"
/db_xref="GI:2642157"
/translation="MKPLTNGDLFKSPTLIKI SALVFTVTVAFYLGKMSDDGYQOLV
FFSSSTSGSIEVSPVSNRNFVLSAIIPNHTQIEIPATIRQPPSVVADTEKV
VEANPPPPPPPPPPPPGPGVFCIVDAGVMSDDFEVGEVSDTVEDWGNQTEIV
EAKSDSKARVRIKFGPCPSRMREYIPLDNTDVIKKLTERGERFERHCPEKGG
GKLVCPKRVKQIPWPKSRDEVNFVNPHTRIVEDKGGONWISROKPKFKPGK
TQPIHGAQYLDQSKWSDITFEKHIRVANDVGGVSPGAYILSRDVTMSVAPKD
VHNGIQFALERGVAAAFATRLLYPSQFDLHCSRKNWTRDDGLLLSINR
MLRAGYFAWAQVYVKEFALEEQWTEMLNTISLCWLVKKEGVVAIWQKPFND
YLSRDSYIPLCDESDPDNWTYTNLKPCI SRIPEKGYGNVPLWPARLHTPPDRLO
TIFRSDYIARKELFKAESKYNEIIGVYRALKWKKMLNVLDMRAGFGGFAAALND
HKLDCWLVSPVPGNTLPVLYDRGLLGVMHWCPEPDYTPRTYDFLHAGSLFIER
KRCENSTILLEDRILRPGRAYIRDSIDVMBEIQEIKKANGWHTSLRDTSEGPHAS
RLITCKRLIRA"
repeat_region 24452. .24497
/rpt_family=" (GGA)n"
repeat_region 25043. .25065
/rpt_family=" (A)n"
repeat_region 25149. .25176
/rpt_family="AT-rich"
gene 25623. .28641
/gene="At2g39760"
/note="synonym: T517.6; predicted by gensean; similar to
SP P34568 YNVS_CABEL; supported by full length cDNA:
Ceres:8256"
mRNA join(25623. .26134,26620. .26968,27055. .27331,28305. .28641)
/gene="At2g39760"
CDS join(25746. .26134,26620. .26968,27055. .27331,28305. .28516)
Query Match 25.8%; Score 261.2; DB 8; Length 92624;
Best Local Similarity 98.9%; Pred. No. 1.1e-55;
Matches 263; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 744 AGATGGTGAAGCTATGGTCTGAAGATGCCACAGAAAGAGAGCTCTTGAATCGGAA 803
DB 35997 AGATGGTGAAGCTATGGTCTGAAGATGCCACAGAAAGAGAGCTCTTGAATCGGAA 36056
QY 804 AGATGGGAGCTTTGTTTCTTCTCTGAGAGACATACATACATGCTCTGATGTTGTAACCT 863

Db 36057 AGATGGGAGCTTTGTTTCTTCTCTGAGAGACATACATGCTCTGATGTTGTAACCT 36116
QY 864 TACTACCAAAAACCTATAAAGATTGGCTTAATTCCTTCTATTGGATATGTCATCATCATAC 923
Db 36117 TACTACCAAAAACCTATAAAGATTGGCTTAATTCCTTCTATTGGATATGTCATCATCATAC 36176
QY 924 TGGTAATCAAGTCTCTTCTTAATAATGTAGAAGATCAGAAAATCCATAAGAAATATCA 983
Db 36177 TGGTAATCAAGTCTCTTCTTAATAATGTAGAAGATCAGAAAATCCATAAGAAATATCA 36236
QY 984 ACATTGAGTCTTATGGTAAAAAAA 1009
Db 36237 ACATTGAGTCTTATGGTAAATGAA 36262
RESULT 5
ATH275979
LOCUS
DEFINITION
Arabidopsis thaliana gene for GDP-mannose pyrophosphorylase, gene
for vacuolar ribonuclease and gene for proline byosynthesis
coding enzyme.
ACCESSION
VERSION
KEYWORDS
GDP-mannose pyrophosphorylase; proline byosynthesis codifying
enzyme; vacuolar ribonuclease.
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS
TITLE
Genes responding to phosphate starvation placed together in
Arabidopsis genome
JOURNAL
REFERENCE
2 (bases 1 to 9423)
AUTHORS
TITLE
Direct Submission
JOURNAL
Submitted (01-MAR-2000) Avila C., Biologia Molecular y Bioquimica,
Facultad de Ciencias, E-29071, Malaga, SPAIN
FEATURES
Location/Qualifiers
1. .9423
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
mRNA join(1405. .1515,1597. .1659,1748. .1856,1883. .2013,
2099. .2770)
CDS join(1405. .1515,1597. .1659,1748. .1856,1883. .2013,
2099. .2770)
/codon_start=1
/product="GDP-mannose pyrophosphorylase"
/protein_id="CAC35355.1"
/db_xref="GI:13509287"
/translation="MKALLVGGFGLRPLTILSPKPLVDFANKPMLHQLKALKAV
GVEVLAINYQPEVMNLFKDPETKLEIKITCSQETPLGTAGPLALARDKLDGSG
EPFVLNSDVISEYPLKEILFEFKSHGGEASINVKVDESKYGVVVWESETRVEKF
VERPKLVGNKINAGIYLLNPSVLDTKIELRPTLSEKTEFFKIAAAQHYAAWLPFWN
DIGQPRDITGLRLYLDLSLRKSPAKLTSGPHVGNVLVDETATIGGCLIGDPAIG
PGTIVSGVRLSRCTVMRGVRIKKHACISSIIIGHSTVGQWARIENMTILGEDVHVS
DELYNSGVVLPHKEIKSNILKPEIVM"
1405. .1515
/number=1
1516. .1596
/number=1
1597. .1659
/number=2
1660. .1747
/number=2
1748. .1856
/number=3
1857. .1882
/number=3
exon
intron
exon
intron
exon
intron

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Nicotiana.

1 Hayaashi,T., Kobayashi,D., Kariu,T., Tahara,M., Hada,K., Kozuma,Y. and Kimura,M.
Genomic Cloning of Ribonucleases in Nicotiana glutinosa Leaves, as Induced in Response to Wounding or to TMV-infection, and Characterization of Their Promoters
Unpublished
2 (bases 1 to 4210)
Hayaashi,T. and Kimura,M.
Direct Submission
Submitted (10-JUN-2003) Takeshi Hayaashi, Kyushu university, Laboratory of Biochemistry, Department of Bioscience and Biotechnology, Faculty of Agriculture; Hakozaaki 6-10-1, Higashi-ku, Fukuoka, Fukuoka 812-8581, Japan
(E-mail:takeshi@agr.kyushu-u.ac.jp, Tel:81-92-642-4215(ex.4215), Fax:81-92-642-2854)
Location/Qualifiers
1. 4210
/organism="Nicotiana glutinosa"
/mol_type="genomic DNA"
/db_xref="taxon:35889"
join(1..248,1322..1350,1443..1516,2308..2412,3238..3330,3926..4210)
/gene="ngr2"
join(1..248,1322..1350,1443..1516,2308..2412,3238..3330,3926..4210)
/gene="ngr2"
/codon_start=1
/product="ribonuclease NGR2"
/protein_id="BAC77612.1"
/db_xref="GI:31621000"
/translation="MTPLPVHAI VNLRLMLFLIIGACILVGVNGDVGELLRRGRG GAGGYOREFDYFKLSIQWPTCYRRTRRCSSNACSRNSRPVFTIHLWTEYNDG TWACCGKAPDEREISTLLEPKYWPSSLCSQSPRSCHERKQDPMWAEWEKJCTAY PVWDEYFFLTINIFYKNVTVEFAGTVPSDSKYPILGGIISIEIAFHATPDL TCSGDALAEELRICFYKNFEPFRDCARDTSALSRGCPQYVSLFAHGSWDEIRISNTTAA S"
intron 2413..3237
/gene="ngr2"
/contig_id="5"
/contig_size="1507"
/contig_start="1"
/contig_end="1507"
BASE COUNT 1175 A 733 C 795 G 1507 T
ORIGIN
Query Match 10.1%; Score 102.2; DB 8; Length 4210;
Best Local Similarity 62.9%; Pred. No. 3.7e-15;
Matches 158; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 498 GGATGTCCTTTATCAAGCTGCTATGTTGCTTCCACAGTCGAAAGTATCTCTAGAGG 557
Db 3925 GGAAGTCTGTGTTGAAGCTGGATGATGACCATCAGATTCGAAAGTATCCATTAGGAGG 3984
QY 558 TATCGTAACAGCCATTGAGATTCATATCACCCCTCAAGTGGTTTGCAAAAGAGA 617
Db 3985 CATCATTTTCATCAATTGAAACGGTTTTCATGCAACCCAGAGTTACATGCTCAGGCGA 4044
QY 618 TCGAATCGATGAATAATGATATGCTTCTATATAAGATTTTAAAGCCAGGAGCTGTGTTGG 677
Db 4045 TGCCTCGAGGAACCTCGTATATGCTTCTATAAAGATTTTCAGCCTCGTGTGTTGCACG 4104
QY 678 TTCACAGATTGCACNCTGAAAGTCATGCCCCCAAGTACCTAGTTTCCCGGAATACAC 737
Db 4105 CGATCTAGCGCCTTATCAAGAGGTCATGTCCTCAATATGTCAGCTTCCAGCCCATCG 4164
QY 738 GCCATTAGATG 748
Db 4165 ATCATGGGATG 4175
RESULT 8
AP006430
LOCUS

DEFINITION	Lotus japonicus genomic DNA, chromosome 5, clone:ljT33L13, TM0327, complete sequence.		
ACCESSION	AP006430		
VERSION	AP006430.1 GI:31581061		
KEYWORDS	HTG.		
SOURCE	Lotus japonicus		
ORGANISM	Lotus japonicus		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.		
REFERENCE	1		
AUTHORS	Asamizu,E., Kato,T., Sato,S., Nakamura,Y., Kaneko,T. and Tabata,S.		
TITLE	Structural Analysis of a Lotus japonicus Genome. IV. Sequence Features and Mapping of seventy-three TAC clones which cover the 7.5 Mb Regions of the Genome		
JOURNAL	DNA Res. (2003) in press		
REFERENCE	2 (bases 1 to 105384)		
AUTHORS	Sato,S.		
TITLE	Direct Submission		
JOURNAL	Submitted (07-MAY-2003) Shusei Sato, Kazusa DNA Research Institute, Department of Plant Gene Research; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 232-0818, Japan (E-mail:ssato@kazusa.or.jp, URL:http://www.kazusa.or.jp/, Tel:81-438-52-3935(ex.2337), Fax:81-438-52-3934)		
FEATURES	Location/Qualifiers		
source	1..105384		
	/organism="Lotus japonicus"		
	/mol_type="Genomic DNA"		
	/db_xref="taxon:34305"		
	/chromosomes="5"		
	/clone="ljT33L13"		
	/clone_lib="ljT library"		
	/note="TAC clone:TM0327"		
BASE COUNT	34361 a 18448 c 18695 g 33880 t		
ORIGIN			
Query Match	8.2%; Score 82.6; DB 8; Length 105384;		
Best Local Similarity	68.0%; Pred.No.3.5e-10;		
Matches	115; Conservative 0; Mismatches 54; Indels 0; Gaps 0;		
Qy	492 TGTCAAGCTGTCCTTTATCAAGCTGGCTATGCTTCCAACTGAAAAGTATCCCTCT 551		
Db	58021 TTTCAGACTCTCTTAATGAGCGTGGTATGTTCCATCTACACAGAAAAGTTCCCT 58080		
Qy	552 AGGAGTATCGTAACAGCCATTGAGAATGCATTTCATATCACCCCTGAAGTGGTTTGCAA 611		
Db	58081 TCGAGGCACTGTATCTGCCATTAGAAATGCTTTCCATATGCTCCCTTAATAATTTGCTC 58140		
Qy	612 AAGAGATGCAATCGATGCAATACGATATGCTTCTTATAAGATTTTAAG 660		
Db	58141 AAAAGTTCTGTGGAGAGCTCGGCTATGCTTCTATAGGCTTCAAG 58189		
RESULT 9			
AMA489249			
LOCUS	4937 bp DNA linear PLN 24-OCT-2002		
DEFINITION	Antirrhinum majus X Antirrhinum hispanicum sl28 gene for S-like RNaase, exons 1-8.		
ACCESSION	AJ489249		
VERSION	AJ489249.1 GI:21615404		
KEYWORDS	S-like RNaase; sl28 gene.		
SOURCE	Antirrhinum majus x Antirrhinum hispanicum		
ORGANISM	Antirrhinum majus x Antirrhinum hispanicum		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; lamials; Lamiales; Antirrhinaceae; Antirrhineae; Antirrhinum.		
REFERENCE	1		
AUTHORS	Liang, L., Lai, Z., Ma, W., Zhang, Y. and Xue, Y.		
TITLE	ASL28, a senescence- and phosphate starvation-induced S-like RNaase gene in Antirrhinum		
JOURNAL	Biochim. Biophys. Acta 1579 (1), 64-71 (2002)		

complement(24618. .28304)
 /gene="At2g16560"
 /note="synonym: F1P15.6; pseudogene, reverse transcriptase"
 /pseudo
 28430 .28467
 /rpt_family="AT_rich"
 repeat_region
 complement(28430. .28467)
 /rpt_family="AT_rich"
 28602 .28640
 /rpt_family="AT_rich"
 repeat_region

Query Match 5.5%; Score 55.2; DB 8; Length 56284;
 Best Local Similarity 73.7%; Pred.No.0.0034;
 Matches 112; Conservative 0; Mismatches 33; Indels 7; Gaps 3;

QY 27 TTTATGCTCTTCCTCTCTGTTGCTGATCGCGGAGCATTTCCGCGAGACGTCATCGA 86
 Db 7060 TGTGTTTTTTCCTCTCTGTTGCTGATTTATCACCAGACGATT--GCGGAGATGCGTCAA 7117

QY 87 ATCAATCATGCTCAGAGGAGTTCGATATTTGCTCTATCTCTTCAATGGCTGGAAC 146
 Db 7118 AGTCAATGCTCTCCCATAAATTCGATTATTTGCT--CTGGCTCTTCAATGACATGAAC 7175

QY 147 CTTATGCGGTGGAATCGCATTTGCTTCTCCA 178
 Db 7176 CTAATGCT--TGAATCACCATTGATGCTCAA 7204

RESULT 11
 AP003349/c
 LOCUS
 DEFINITION
 Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC clone: P0674H09.
 AP003349 BAO000010
 VERSION
 AP003349.4 GI:22093582
 KEYWORDS
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
 1
 Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T., Katayose, Y., Wu, J., Nilmura, Y., Cheng, Z., Nagamura, Y., Antonio, B.A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K., Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K., Hamada, M., Harada, C., Hijishita, S., Honda, M., Ichiikawa, X., Itonuma, A., Iijima, M., Ikeda, M., Ikono, M., Itoh, S., Itoh, T., Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S., Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T., Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M., Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T., Shomura, A., Song, J., Takazaki, Y., Terasawa, K., Tsuji, K., Waki, K., Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H., Iwana, H., Endo, T., Ito, H., Hahn, J.H., Kim, H.I., Eun, M.Y., Yano, M., Jiang, J. and Gojobori, T.
 The genome sequence and structure of rice chromosome 1
 Nature 420 (6913), 312-316 (2002)

TITLE
 JOURNAL
 MEDLINE
 22337376
 PUBMED
 12447438
 REFERENCE
 2 (bases 1 to 142955)
 Sasaki, T., Matsumoto, T. and Yamamoto, K.
 Direct Submission
 Submitted (28-FEB-2001) Takuji Sasaki, National Institute of Agrobiological Sciences, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsukuba.ibaraki.305-8602.jp, URL: http://rgp.dna.affrc.go.jp/, Tel: 81-298-38-7441, Fax: 81-298-38-7468)
 On Aug 2, 2002 this sequence version replaced gi:20975423.
 GENES were predicted from the integrated results of the following: GENSCAN1.0, RAST2.0, BLAST2.0 as well as SplicePredictor

October 1999 version). The genomic sequence was searched against NCBI Nonredundant Protein database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologs of the coding regions were searched against NCBI Nonredundant Protein database with BLAST2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone ID.
 A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.
 The orientation of the sequence is from 5' to 3' of the PAC clone. This sequence of P0674H09 clone has an overlap with P0696G06 clone (DDBJ: AP003316) at the position 1 to 51,418 of 5' end and an overlap with P0506A10 clone (DDBJ: AP003418) at the position 55,654 to 142,955 of 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at http://rgp.dna.affrc.go.jp/GenomeSeq.html.
 Location/Qualifiers
 1. 142955
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="1"
 /clone="P0674H09"
 5794 . 6351
 /gene="P0674H09.1"
 5794 . 6351
 /gene="P0674H09.1"
 /note="hypothetical protein"
 /codon_start=1
 /protein_id="BAC06879.1"
 /db_xref="GI:22093583"
 /translation="MGKRGPCGSRHGVARCLAWRLARRHGRWRNCGEGKLAAM EAVGKAVADVGDIDCFLLPNGLLLALLPSPSEFCRKLPIKNGYVGRKVKPHVITITISERDYRHFYRQLPITPTTVGANRKRKPESESVSPPLGRTRWCLV RGAGRLLAARLSRVIAWAESR"
 join(8132. .8265,8497. .9025)
 /gene="P0674H09.2"
 join(8132. .8265,8497. .9025)
 /gene="P0674H09.2"
 /note="hypothetical protein"
 /codon_start=1
 /protein_id="BAC06880.1"
 /db_xref="GI:22093584"
 /translation="MAASARRDSGLCPKPKRVRPESTSRVLARAGSDTRDKR RAVACQGEASVRRKAGNGGGERVEFCFPGGCGWERALPDGRTAAOLGHR WLAQPSATTAARQGAARAAATARAARAGTVCDAKPAFGSGSPGRWEGERE NERREKALGLAWRLTTRVQMRGARGLREREREMGREINRPTPEGGKLDFFR GI"
 join(14423. .14497,14878. .15690,15782. .15901,16516. .17199)
 /gene="P0674H09.3"
 join(14423. .14497,14878. .15690,15782. .15901,16516. .17199)
 /gene="P0674H09.3"
 /note="contains EST D39962(S1657)
 similar to Arabidopsis thaliana chromosome 3, At3g59690 unknown protein"
 /codon_start=1
 /protein_id="BAC06881.1"
 /db_xref="GI:22093585"
 /translation="MGKRGWITALKKFASTGPKDKNTGNQLVAQYSHQHSRGGGA AARDKRWGQGRSGHAEFSPAGALINILYREPSSELEKILVDHMEQQRQVRAHQV ITPKAKTATPAAAAAAPPAPPVPSARERQRDRDDKPAIVLPLPSPSPVPIRR FDHREQOQKLQOQOQTTPRWRPQAORRAARPPADPPADRAAAVAQTQAFR GYMAERNYRGLRLRLQGVVGRGSRVQTAHMRQMLRVQSVQSRVRAAMER NRHHAMLRDAARWRAASODGQWEDSLISLRDMDATKTKVEATVKRRLAYAYS HCLLTMAAHAILADLOQRNFWPTPIERHEAAALSRQVEHVHGSGGGG GVALVNTPAKSVSVSSTATTATTAAPRATPRAKVASSVGGGSIQDLSLSCSPF RGATVNTPTMSASAKARQAQMLRQOQAQAQEKPRFSGGLQGSIGSWAKSFF

gene
WKGGAAPPSISSRVATPAASVAGHRPRTSSVSELSVDSVAVMPAGIGRRTFK"
join(20555..21175,21273..21372,21678..21822,22501..22624,
23076..23609)
/gene="P0674H09.4"
CDS
join(20555..21175,21273..21372,21678..21822,22501..22624,
23076..23609)
/gene="P0674H09.4"
/note="hypothetical protein
similar to Arabidopsis thaliana chromosome 1, F5114.2"
/codon_start=1
/protein_id="BAC06882.1"
/db_xref="GI:22093586"
/translation="MSGKLLGVAPHPANNIDADLYDLGOSRMQVDPVVLNRSAE
LESDMADDDVVEASSOPAPAKQSPALMDTIVEVOKLKRKASSGSLAABKDA
LVAGCCCELEGLLEITREVSGRNQFNGLNTNAAIGCLLESGLSSGLKVDIEYER
LKMEGVSSVSSVSLVILQGRWYQSSPDADVLEDESETALWCWENVDLKVILPRM
RGLPSTRTARKHRIITAITYSTLSVLEAPGAEQVNDMRKASLKLKALNLEIGKS
LVRATQKSNIEGAKNTGTAKEMQVKNNDGTIITENVDDSLQKNTNTSEKDT
OKAQVKELKQKEEAKRQKQOEALREOKREKEAVOLKQALOKQASMERFFPKK
KEOKREKEAE"RKQKQOEAEKEQKREKEAVOLKQALOKQASMERFFPKK
KDEKLEKPGKDSGVQTTDPTCNKEVPLVTSIIDSFSQENWALDRLRLISF"
join(23955..24091,24190..24271,24342..24477,24697..24737,
24942..25340,25423..25560,25696..25741,25833..26002)
/gene="P0674H09.5"
CDS
join(23955..24091,24190..24271,24342..24477,24697..24737,
24942..25340,25423..25560,25696..25741,25833..26002)
/gene="P0674H09.5"
/codon_start=1
/product="putative PAS1"
/protein_id="BAC06883.1"
/db_xref="GI:22093587"
/translation="MLPAVELOPHGNTNANPLPTRISIKRLQLQDKSNRPAYGTWRK
KSAVGPRLKMDPLDVEDSDMEWEDPESLSDCEKNDREWEDSDITDBES
EDSFVDPDQNEAKNTVEQKSLVEGQIESLLDDKDEASSPDQCAEVEEF
RALRQKVLNLTGALKRSQPLVISNLTHEKAEILLTAGDLKGRSKIQGLQVLSM
RICFGGATIDPLVIDSSANAEITNQLNVKSSPAAAIPTDILABIVIGSCRDGI
NKVLESHQRPVSSQLKRVREISFEFVNNRWQKVEVLKSLGSSSPASSKPKPS
IATYFSKRCIPPEEALIASPELRKSKTQNVNGDTDIPRIINLLPSSQ"
complement(26516..27016)
/gene="P0674H09.6"
CDS
complement(26516..27016)
/gene="P0674H09.6"
/notes="contains EST C27797(C52876)
unknown protein"
/codon_start=1
/protein_id="BAC06884.1"
/db_xref="GI:22093588"
/translation="MBAQGSHTHRAFLLCNVALLGAASSCFLTSLSLSPGCLL
LLEHALTAVPSAAGSGSTPATPAQWNAHTAGALTAIFQGVALLATRTSDP
LAEQSVTRDDDAVILKMGVGLGTAFVLEWALALAFSLRDLDDGAGDGNKWA
AASYHV"
complement(join(28586..28757,28943..29001,29174..29220,
29897..29927,30140..30202,30393..30488,30572..30655,
30745..30837,31138..31215,31571..31661,31741..31841,
31936..32754,32849..32899))
/gene="P0674H09.7"
CDS
complement(join(28586..28757,28943..29001,29174..29220,
29897..29927,30140..30202,30393..30488,30572..30655,
30745..30837,31138..31215,31571..31661,31741..31841,
31936..32754,32849..32899))
/gene="P0674H09.7"
/note="contains ESTs AU057282(S21302), AU057281(S21302)
similar to Arabidopsis thaliana chromosome 3, Atg59780
unknown protein"
/codon_start=1
/protein_id="BAC06885.1"
/db_xref="GI:22093589"
/translation="MLLPVCSAAPTCSPLCVPTSSVAHVVGAPALRGVRVVEALSRSY
ADPLVSPDRPLADSSILSPVSPDDIVGLAAAEVPSADATAAGDALCCGVE
PATLVASDLAATTADAATDAERALLDAPVTTTPADASDAEAFALIDSLGKI
FOADELUTEGDLTSAVLSLAWKRSVQGVGLTASVDATKQAGVDTASGAL
QDKVAGATVADVILRAIVAEDSLGSAATISGYGGTSSLPNPNVLDLNSSEK
ASLVLPKIGGALQVYVVEGIEKNVGLDFSDPIVQLVALLGSLTIGLTWLPAYGG

gene
YSGDLSPESTLELRNDKADLRVKDGIPLDRAARSKYATAASPEIKGSRVRLKSG
RDVDDALLAVIRNLKLVGDSKVIINDANGSRKATRLKLGVRPVLVKGQFS
WAKLRVKEKLPETALVINEDEEILEDIKPTFTLVGLSLGSAAYALLEWETTL
QYGVLSGLITVYLFSTYEGSEDLQULKVGAEAFSWAAKLEPSPKIGLATSPTTA
VODVLOAAKHESQPSDABADSLVSEA"
complement(join(34221..34296,34561..34647,35201..35291))
/gene="P0674H09.8"
CDS
complement(join(34221..34296,34561..34647,35201..35291))
/gene="P0674H09.8"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAC06886.1"
Query Match 5.4%; Score 54.6; DB 8; Length 142955;
Best Local Similarity 53.5%; Pred. No. 0.0047;
Matches 114; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 269 GTTCTTATTCGATCTGACCTTTAAAGAGAAGGAGATTTCACCGTTGATGATCGTCTTGAGA 328
Db 63672 GTGTGTAATTATCTGGTTTACCTTTGTGACAGATCTCGAGGTTCGAAGACGATATCTCGAGG 63613
QY 329 AGTACTGCGCTAGTCTCGATTGTGTTCTCCATCATCATGCAATGGTGGGAAGGGTTCAT 388
Db 63612 AGTACTGCGCTGCTCTTGTACTGCGGCTCTTTCTCAACCTGCTTTGGTGGGAAAGGCCAT 63553
QY 389 TTTGGGGCCACAGTGGGAGAAACATGGGAGCTTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 448
Db 63552 TTTGGGTCCATGAGGTAGACATTGGTATCTAAATAGATATGTGATTGTGATGAGGAGAAA 63453
QY 449 ATAATTACTTCTTACCAACACTTAATCTCTACT 481
Db 63452 TTACTGACATCTTACCCTTTTTCCTCTCTGACT 63460
AP003418 145265 bp DNA linear PLN 29-WAR-2003
Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1,
PAC clone:P0506A10.
AP003418 BA000010
AP003418.2 GI:15216346
Oryza sativa (japonica cultivar-group)
Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartioideae; Oryzeae; Oryza.

REFERENCE 1
Sasaki, T., Matsumoto, T., Yamamoto, K., Sakata, K., Baba, T.,
Katayose, Y., Wu, J., Niimura, Y., Cheng, Z., Nagamura, Y.,
Antonio, B. A., Kanamori, H., Hosokawa, S., Masukawa, M., Arikawa, K.,
Chiden, Y., Hayashi, M., Okamoto, M., Ando, T., Aoki, H., Arita, K.,
Hanada, M., Harada, C., Hijishita, S., Honda, M., Ichikawa, Y.,
Idonuma, A., Iijima, M., Ikeda, M., Ikono, M., Itoh, S., Itoh, T.,
Itoh, Y., Itoh, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Katagiri, S.,
Kikuta, A., Kobayashi, N., Kono, I., Machita, K., Maehara, T.,
Mizuno, H., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nakashima, M.,
Nakama, Y., Nakamichi, Y., Nakamura, M., Namiki, N., Negishi, M.,
Ohta, I., Ono, N., Saji, S., Sakai, K., Shibata, M., Shimokawa, T.,
Shomura, A., Song, J., Takaraki, Y., Terasawa, K., Tsuji, K., Waki, K.,
Yamagata, H., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K.,
Zhong, H., Iwama, H., Endo, T., Ito, H., Hahn, J. H., Kim, H. I., Eun, M. Y.,
Yano, M., Jiang, J. and Gojobori, T.
The genome sequence and structure of rice chromosome 1
Nature 420 (6913), 312-316 (2002)
22337376
PUBMED 12447438
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE 2 (bases 1 to 145265)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Direct Submission
Submitted (14-MAR-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan

(E-mail:tsasakienias.affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>, Tel:81-298-38-7441, Fax:81-298-38-7468)
 On Aug 20, 2001 this sequence version replaced gi:13366121.
 Genes were predicted from the integrated results of the following:
 GENSCAN1.0, BLASTX2.0, BLASTX2.0 as well as SplicePredictor
 (October 1998 version). The genomic sequence was searched against
 NCBI NonRedundant Protein database, nr
 (<ftp://ncbi.nlm.nih.gov/blast/db/>) and the cDNA sequence database at
 RGP. Protein homologies of the coding regions were searched against
 NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
 the identified cDNA sequences using BLASTN 2.0 with the
 corresponding DBJ accession no. and RGP clone ID.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with EST homology (covering
 almost the entire length of partial sequence) is classified as an
 'unknown' protein. A gene predicted with a gene prediction program
 is classified as a 'hypothetical' protein.
 The orientation of the sequence is from SP6 to T7 of the PAC clone.
 Detailed information on overlap and assembly quality together with
 annotation of this entry is available at
<http://rgp.dna.affrc.go.jp/genomeSeq.html>.

FEATURES

source
 1. 145265
 Location/Qualifiers
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="1"
 /clone="P0506A10"
 complement(join(2525..2595,2686..2824,3342..3503,
 4105..4206,4353..4426,4567..4595,4746..4969))
 /gene="P0506A10.1"
 complement(join(2525..2595,2686..2824,3342..3503,
 4105..4206,4353..4426,4567..4595,4746..4969))
 /gene="P0506A10.1"
 /note="contains ESTs C96756(C10648),D22272(C10648)"
 /codon_start=1
 /product="putative S-like ribonuclease RNS2"
 /protein_id="BAB86068.1"
 /db_xref="GI:19386686"
 /translation="MATARKAPRWAIWVVVVVAAADLFGAAVASASARHAAVGKKO
 REFDYFALALOWPTGACRHCACNGCCRSPLQTFTHGLWPDYDGTWPACCRH
 TSDMDKILPTLEKYPWLSYSSSTCFSGKPFWAHEMLAGGHVHNGKQYA
 LTDVIDAKAFAGSQVCKGSEVELKCFDKKTKTILLNSVYFQLDCLTT
 TATNENYKKKCYPRYITLPTDPIVHANSIRITIVSESVGYLTS"
 complement(join(5975..6058,6439..6510,6901..7062,
 7248..7304,7886..7987,8137..8210,8302..8330,8428..8648))
 /gene="P0506A10.2"
 complement(join(5975..6058,6439..6510,6901..7062,
 7248..7304,7886..7987,8137..8210,8302..8330,8428..8648))
 /gene="P0506A10.2"
 /note="contains ESTs AU092918(C60895),C28387(C60895)"
 /codon_start=1
 /product="putative S-like ribonuclease RNS2"
 /protein_id="BAB86069.1"
 /db_xref="GI:19386687"
 /translation="MEERQRLVTSALCLAVAMSGALLPPRAAGVGRKRWAA
 GFDYVIALQWPGVTCQTHCCSSNGCCSHLKFETIHLMPQSYGGWPCCRPT
 TDKNKISRLKLTILEYWPGLYCGSTFCGKRPFWHEDEYDFSTALYLYKYNV
 TKALKAHIVPRGKRYLVGHVSAIEFSGAMPISIVCKNGSVQELRLCFHKDYQLRV
 RTLEETALDTSVLSVHINHTLAILOKESVTSKTLSTNSHSLS"
 join(11526..11584,11964..13698)
 /gene="P0506A10.3"
 join(11526..11584,11964..13698)
 /gene="P0506A10.3"
 /note="hypothetical protein
 similar to Arabidopsis thaliana chromosome 3, At3g02650"
 /codon_start=1
 /protein_id="BAB86070.1"
 /db_xref="GI:19386688"
 /translation="MFHGNADNSIEDLSLSRNYHAAAAAADAEEEMWRSRGARLL
 LBSSHTRLPPPPSPRLALTRAPPPRLSPFLSSSPRLPDAGGGAASASSDPTDA
 GEGCLASDPTEAGEDNLASLWEEDAGDADDVAADSDPAADVDEVLAARVGAUVE
 STEPDQPSALADAMVDFNEPLIALILVSAASCCSKLLRLPKSGKNNPAKSLAN
 LEIVDKIADSDIDVDVLLWDLWKEWAGVSTVQVNLKLLAMFWLKEKSKAL
 EAFDMFSELCGAPGDSYLAIOAGKAGMVDAAWRCGMISSGCFPGDEKVGDIUM
 FFKCKKYKRAHSVITAKEKQVAPTSLTFLVSALARNDENTINALELSEYQGES
 PKGQSYAAVTHGLCTNNVEDAKVLMVNLGAPGSAVFNFVITALLSKNGMED
 KGLMRVMENQVICPDITYTYSVMGKAGMMDEAHALRDARLHPLRSVYTHL
 IRGCKWEERPEKALECLKEMKEDGPMNDENYKLIQSLCLKALDWRTAENLLKEMED
 GGJCLKGTTRSLIAAVKLEMDLSKQSEA"
 join(16095..16181,16234..16414,17357..17415,17616..17691,
 17836..17910,17987..18074,18241..18466,18669..18784,
 19152..19363,19862..19896,19985..20087,20625..20733,
 20818..21016)
 /gene="P0506A10.4"
 join(16095..16181,16234..16414,17357..17415,17616..17691,
 17836..17910,17987..18074,18241..18466,18669..18784,
 19152..19363,19862..19896,19985..20087,20625..20733,
 20818..21016)
 /gene="P0506A10.4"
 /note="contains ESTs D41563(S4135),C98908(E3241)"
 /codon_start=1
 /product="putative beta-glucosidase"
 /protein_id="BAB86071.1"
 /db_xref="GI:19386689"
 /translation="MWGHLFGRQSLRGSHRLVPPVPTVAAHAETTRGARLTCTATR
 RARCGDRLLAAAGNGEHGARGGDRPLRRLRRLRRLRRLRRLRRLRRLRRLRRLR
 DHRYKEDIEMASLQFRAYRFSISWPIFPDGLKGVNEQGVAFVYDNLINIEAG
 IEPTATLYHWDLPNLQOTVGGWLSKXIVFYFALYACAFNDFGRVKKHTINIEPG
 TAVNGYIGHPAPGCGEGTARCYLAHAYQILAHAAADVRRKFAVQCGEGLVVD
 CEWAPESEKTEQDVAERLDLPOLGWLDPYVFGDVPESMRORLDLPTFSKDKKE
 FIRNKIDFVGINHTSFIAHQDDPEDIYRVVQOVERIEKWTGKIGRAASEWL
 VFWGLRKLNLNAAKRYGNPVITYENGMEDEDDQSATLQVLDNTFVGYFKGLAS
 VQAIKDGDADVRGYFAWSLDFNFWANGYTRFGIVTVYDINLSRHPKASARWFSR
 LKGDAAENKADMN"
 complement(join(2252..22360,22835..22882,23187..23197))
 /gene="P0506A10.5"
 complement(join(2252..22360,22835..22882,23187..23197))
 /gene="P0506A10.5"
 /note="hypothetical protein"
 /codon_start=1
 /protein_id="BAB86072.1"
 /db_xref="GI:19386690"
 /translation="MNDNRLIGDMSLTGYKDVSEVIPDRAISLRCSEKSGGLCMNQ
 SKGSKQKDAEE"
 complement(join(25382..26209,26300..26547,27139..27344,
 27434..29034))
 /gene="P0506A10.6"
 complement(join(25382..26209,26300..26547,27139..27344,
 27434..29034))
 /gene="P0506A10.6"
 /note="contains EST D24760(R2517)"
 /codon_start=1
 /product="putative formin-like protein AHF1"
 /protein_id="BAB86073.1"
 /db_xref="GI:19386691"
 /translation="MPSLRRELVVWVGVVGVGGSDGLEVAVARQLHPFPDQ
 SSSPTPAPPQPPFPFALVPAIPLLLTAVLGLSIAPFFTHRRGNRAAGGGGCV
 GGGGGKGGKSSASLPAIPLLLTAVLGLSIAPFFTHRRGNRAAGGGGCV
 GGGDAKTLHPERASIFAPDEFGSGGAAAPAAADVRYGNAGIGRMDEKSETTSS
 GGBAERSTGPELRPLPPLARCGGAGASPGSGVGFASPGSPGDEEYFVSPOGSSK
 MSTSHRTLAARAAVADKRSKPSGSIYSTSPSPSPGATWSPAPSPPLSPLS
 QSGRSVKSRSVSTFGQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP
 ALRSTTTDTTTPRNFVQPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP
 PLHMDKPPASDVRVMVQKLSQSFQVNEEMETIFICNPANSAPPKPEPTRRPLK
 ETKDRKFEKTSPIKLPASKFKAVLDIPAFKRVAMLYIANFSEVNVYKKSFE
 EITKACBELRNSRLFKULEAVLKTKGNMNVGTNRGDAAKLTLLKLVKVGTDGK
 TTLHFVQVETIRTEGSHLSASNSTPTQANPLRDELECKKGLGLQVAVGLNLSV
 KKAAMDSDVLSYVSKLAGIEKITEVLRNEEVKREDARFHDHSMQKTLKRAADD
 IIRVQAQSVALSIVKEITEYFHGDSAKEAHPFIFMVVDFJUSVDQVCKEVRIN
 DRTIASSVRHPVPVNMWPMQLFPRHAIKAGISDDESSATSSASP"

d'Interet Public, Centre National de Sequencage - GENOSCOPE; 2 rue Gaston Cremieux, BP191, 91006 Evry Cedex, France;
http://www.genoscope.cns.fr
Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/..

FEATURES

source
1. .103098
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
448. .2895
/gene="TSC2_10"
complement(join(448. .1062,1145. .1610,1689. .1759, 1855. .2375,2463. .2663,2799. .2895))
/gene="TSC2_10"
complement(join(448. .1062,1145. .1610,1689. .1759, 1855. .2375,2463. .2663,2799. .2895))
/gene="TSC2_10"
/note="similarity to hypothetical protein At2g15420 - Arabidopsis thaliana, EMBL:AC006920"
/codon_start=1
/product="putative protein"
/protein_id="CAB81572.1"
/db_xref="GI:7263606"
/translation="MDCQWCDLNGYTNQDAYRGINDGLPVDVDFGGQMSREVDDEDDI VYTPHRISIVGLRSLRTPGLDSVVVPPDRDGSGLCTRKSLTLIDGSIIPDDAG LRFPLDRLVYCLRRRIATFSQLALAVCNAGLMIGADREVEVDAGLLEATTAA PKDNGLCQVARNPHKLVCKSKVDRWLYFFVZLSISVSDLDLVLFASGICF LVDFSLARVTVNVAFLPRULVFAHDFQRGTFGGISASLAARTRKPLHWPDIHR QQRWEKLAISFRFTLRSGSEIETNOQITNEDRATQSDARVKKASGSDSDRAM RLDDARAKIASNVTEGAIFFREPLSVHAPSKVSHGSSRDSFPVHERGSTQDI REVDFSFNTDRKGYFPFSDRGYCGFPAKRVNGLSRNPVPSPTAKLYQWACKE VQSIATGNALYNFLFEOQAREAEEMVAEAKKIAETAAATRLRLNLSKRGTEGEK QILESLKXSDSRGAEQSRSLRSSRDDALRTTRSRSRKDDIALYKARIDLRLKH ILDKRVADSPFLFMOQVAGSRLLKLLITNGLAVPGTMCGLKVDKXKGVVDLAI TDLDDDLVAFPEVEDERAMTTAGEARIGST"
/gene="TSC2_10"
/number=1
complement(1063. .1144)
/gene="TSC2_10"
/number=1
complement(1145. .1610)
/gene="TSC2_10"
/number=2
complement(1611. .1688)
/gene="TSC2_10"
/number=2
complement(1689. .1759)
/gene="TSC2_10"
/number=3
complement(1760. .1854)
/gene="TSC2_10"
/number=3
complement(1855. .2375)
/gene="TSC2_10"
/number=4
complement(2376. .2462)
/gene="TSC2_10"
/number=4
complement(2463. .2663)
/gene="TSC2_10"
/number=5
complement(2664. .2798)
/gene="TSC2_10"
/number=5
complement(2799. .2895)
/gene="TSC2_10"
/number=6
6611. .8903
misc_feature

/note="homology to fragments of reverse transcriptase"
9985. .11518
/gene="TSC2_20"
complement(9986. .11518)
/gene="TSC2_20"
complement(9986. .11518)
/gene="TSC2_20"
/note="similarity to retrotransposon finger protein (clone DW15) - Arabidopsis thaliana, PIR:S65811"
/codon_start=1
/product="putative protein"
/protein_id="CAB81573.1"
/db_xref="GI:7263607"
/translation="MEPLVASQPPVTGTEKGLQLISLVSVVHKQIDGPEKIVESHQK EDGGSNSDSLVLGSGSEATSEASRGTSHTVTKSWLQAAQKSFTRQKPEV VEVDQKVVVPREVFLGAPLWEDFLVGNILNSKAPHVGIKIMINKLWRLDKSTL IDVLVNESTVTKFRVNEARHRIILNRMMNIMGLPMITSKMTPYSEDVOPALKSIFL WITLSNVPLTMFTDQKLEFLASVVVGKPVRLHPKTEACESEFEAKIIVEADLTQDLPEK YILSGEBEGELDAVHYSPWLSPHQVCKKWHLRDSCLEKQKSTSTLNNSPVSEK EPOKTVBEARHVSFKVVEELFNNLPQSGSVSALETKVQSNKTDDGMSPLKATTKSP MNSVELKGEVSLISNPYAPLSVQEDIREDEIVTTTSEINLVVQSMVTEKAVEIDF QTSISVGGNRNRCPSKSHPTKQPRGSKTAHKVIPQGNTHLKKSPVRDQGNKIIFPH HH"
complement(9986. .11518)
/gene="TSC2_20"
/number=1
13872. .16433
/gene="TSC2_30"
join(13872. .13930,14081. .14166,14579. .14656,14735. .14956, 15045. .15084,15306. .15482,15566. .15855,15916. .16433)
/gene="TSC2_30"
/note="similarity to putative proteins - Arabidopsis thaliana"
/codon_start=1
/product="putative protein"
/protein_id="CAB81574.1"
/db_xref="GI:7263608"
/translation="MGEVSPYTAIKNLKHKFTSCSRHILFCVFGNKLQAGIKKNIY SNKGYAEENGTNDIDDDSDSYNSESSEDEENLQEDFYWSSQESCSLSLDED VNAEDMAQIPDDIIIEAKQKFDILNLENAFGPKSKPVIIPNPKSAYVDEGDPVTSC DKGKVDHCLPKDRGNMFAIQENYYIMGALKPKGAKAFQOLYIVDTENEVKRYN IMNSNESENGQKKKIMKLLNRVNPVKAFRYDRFNTDAQESFHMRIIAARKTDP RYNLPTASEVAALIPGDHEDMDKLDIVLQETSGKLRKINGRIGIKKATNGHGD GKQKQDVSMRQWPAVRLQERKAEKHLFRSKLQOQIVDAVTMTESNRLYIKNQ PKFESSTREPIQNASNSGNDLENQKVKIPASFTGGPRYMTNNYLDAMAICKHFGF PSLEITFTCNPKPPEIVRFCKERNLKSERQDILCLRL"
13872. .13930
/gene="TSC2_30"
/number=1
13931. .14080
/gene="TSC2_30"
/number=1
14081. .14166
/gene="TSC2_30"
/number=2
14167. .14578
/gene="TSC2_30"
/number=2
14579. .14656
/gene="TSC2_30"
/number=3
14657. .14734
/gene="TSC2_30"
/number=3
14735. .14956
/gene="TSC2_30"
/number=4
14957. .15044
/gene="TSC2_30"
/number=4
15045. .15084
/gene="TSC2_30"
/number=5

```
intron 15085. .15305
       /gene="TSC2_30"
       /number=5
exon 15306. .15482
     /gene="TSC2_30"
     /number=6
intron 15483. .15565
       /gene="TSC2_30"
       /number=6
exon 15566. .15855
     /gene="TSC2_30"
     /number=7
intron 15856. .15915
       /gene="TSC2_30"
       /number=7
exon 15916. .16433
     /gene="TSC2_30"
     /number=8
repeat_unit 18563. .18592
            /note="30 bp AT tandem repeat"
gene 19386. 20376
     /gene="TSC2_40"
     complement(join(19386. .19445,19766. .19822,20269. .20376))
gene 19386. 20376
     /gene="TSC2_40"
     complement(join(19386. .19445,19766. .19822,20269. .20376))
CDS 19386. 20376
    /gene="TSC2_40"
    /note="similarity to hypothetical protein At2g39150 -
    Arabidopsis thaliana, ENBL:AC004697"
    /codon_start=1
    /product="putative protein"
    /protein_id="CAB81575.1"
    /db_xref="GI:7263609"
    /translation="MAISEGTMVEGHCVPDLVELMPKHQHDIPRAMLRIVVHEGRNHE
```

```
Query Match 5.0%; Score 50.8; DB 8; Length 103098;
Best Local Similarity 74.4%; Fred. No. 0.044;
Matches 64; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 36 TCTCCTTCCTCGTTCGTCATCGCGGAGCATTTCCGCGAGACGTCATCGAATCAATCG 95
Db 69994 TTTCTTCTCCTTCGCGTTTATCACCGGAGCAATTGTCGGAGATGTCATCAAGTCATCG 70053

QY 96 ATCTCAGAGGAGTTCGATTATTTTCG 121
Db 70054 CTATAACAGAGAAATTGATTATTTGG 70079
```

Search completed: January 21, 2004, 22:14:34
Job time : 3698.68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 15:50:10 ; Search time 260.957 Seconds
(without alignments)
10468.497 Million cell updates/sec

Title: US-09-936-869-5
Perfect score: 1012
Sequence: 1 atcgatttaagtcattggc.....ttctatggtataaaaaaa 1012

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N Geneseq 19Jun03:*	
1:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:	*			
2:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:	*			
3:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:	*			
4:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:	*			
5:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:	*			
6:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:	*			
7:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:	*			
8:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:	*			
9:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:	*			
10:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:	*			
11:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:	*			
12:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:	*			
13:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:	*			
14:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:	*			
15:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:	*			
16:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:	*			
17:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:	*			
18:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:	*			
19:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:	*			
20:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:	*			
21:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:	*			
22:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:	*			
23:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:	*			
24:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:	*			
25:	/SIDSL1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:	*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1012	100.0	1012	21	AAA99365
2	993.8	96.2	1069	21	AAC37325
3	170.8	16.9	822	15	AAQ71177
4	83.2	8.2	266	25	ABX23625
5	76.6	7.6	229	25	ABX23678
6	63.8	6.3	224	25	ABX26579
7	50.4	5.0	1030	21	AAC34348
8	46.8	4.6	693	24	ABZ14066

9	46.8	4.6	1026	21	AAC47291
c 10	43.6	4.3	2155561	24	ABN71527
c 11	42.6	4.2	47108	24	ABK31510
12	40.8	4.0	19513	24	ABL49339
c 13	40.8	4.0	19513	24	AAS61245
c 14	40.4	4.0	10133	24	ABL32458
c 15	40.2	4.0	50000	24	ABL55643
16	40	4.0	34688	24	ABQ67060
17	39.6	3.9	6729	24	ABQ67153
18	39.4	3.9	1278	19	AAV55303
19	38.4	3.9	1278	23	AAS55528
20	39.4	3.9	1287	25	ABX06962
21	39.4	3.9	2764	19	AAV5241
c 22	39.4	3.9	13121	19	AAV52259
c 23	39.4	3.9	2162598	25	ABS56454
c 24	39.2	3.9	365	22	AAI88155
25	39.2	3.9	381	25	ABX19331
26	39.2	3.9	387	25	ABX18894
c 27	39	3.9	7810	22	AAS45434
c 28	39	3.9	7810	24	ABK28281
c 29	38.4	3.8	5520	24	ABL33812
c 30	38.4	3.8	6106	22	AAS46430
31	38.4	3.8	6106	24	ABK40032
32	38.4	3.8	6106	24	ABL33473
33	38.4	3.8	6123	24	ABL32820
34	38.4	3.8	6265	24	ABL33417
c 35	38.4	3.8	10094	25	ABX77219
c 36	38.4	3.8	116277	20	AAZ20249
c 37	38.4	3.8	910715	20	AAZ20248
c 38	38	3.8	789	20	AAZ42490
c 39	38	3.8	789	24	AAD42494
c 40	38	3.8	1342	21	AAZ6424
41	38	3.8	6591	22	AAS46284
c 42	38	3.8	13643	23	ABL03848
c 43	37.8	3.7	757	24	ABQ67170
c 44	37.8	3.7	831	20	AAZ30136
45	37.8	3.7	1290	22	AAH90726

ALIGNMENTS

RESULT 1
AAA99365
ID AAA99365 standard; DNA; 1012 BP.
XX
AC AAA99365;
XX
DT 22-JAN-2001 (first entry)
XX
DE Plant PrAG1 promoter related gene sequence.
XX
KW Plant promoter; PrAG1; reproductive tissue; transgenic plant; cereal; ds.
XX
OS Arabidopsis thaliana.
XX
PN WO200055172-A1.
XX
PD 21-SEP-2000
XX
PF 17-MAR-2000; 2000WO-NZ000031.
XX
PR 17-MAR-1999; 99NZ-0334715.
XX

PA (CART-) CARTER HOLT HARVEY LTD.
PA (TASM-) TASMAN BIOTECHNOLOGY LTD.
PA (UNMT) UNIV MICHIGAN TECHNOLOGICAL.
PI Podila GK, Liu J, Karnosky DP;
DR WPI; 2000-59442/56.
DR P-PSDB; AAB26796.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 21-JUN-1999; 99US-0139839.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140323.
PR 23-JUN-1999; 99US-0140334.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140931.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 99US-0162142.

```
Query Match      96.2%; Score 973.8; DB 21; Length 1069;
Best Local Similarity 99.0%; Pred. No. 5.8e-262;
Matches 1001: Conservative 0; Mismatches 7; Indels 3;
```

QY	1	ATCGAATTAAGTCAATGGCGTCAAGTTTATGTCTTCTCTCTCTCGTTCGCGTATCGCC	60
DB	19	ATCGAATTAAGTCAATGGCGTCAAGTTTATGTCTTCTCTCTCGTTCGCGTATCGCC	78
QY	61	GGAGCATTTGCGGAGACGTATCGAACTCAATCGATCTCAGAGGAGTTGATATTTC	120
DB	79	GGAGCATTTGCGGAGACGTATCGAACTCAATCGATCTCAGAGGAGTTGATATTTC	138
QY	121	GCTCTATCTCTTAATGGCCCTGGAACCTATTCGCGTGGAACTCGGCATTTGTTCCTCAAA	180
DB	139	GCTCTATCTCTTAATGGCCCTGGAACCTATTCGCGTGGAACTCGGCATTTGTTCCTCAAA	198
QY	181	AACGGTCTGTCAGAGGCTCGGATGCTCCAACTCAATTCACAATTCATGGGTATGGCT	240
DB	199	AACGGTCTGTCAGAGGCTCGGATGCTCCAACTCAATTCACAATTCATGGGTATGGCT	258
QY	241	GACTATACGATGTTGCTGGCCCTCATGTTGTTATCGATCTGACTTTAAAGAGAAGAG	300
DB	259	GACTATACGATGTTGCTGGCCCTCATGTTGTTATCGATCTGACTTTAAAGAGAAGAG	318
QY	301	ATTTCAACGGTTGATGGATGGCTTGAGAAGTACTGCGCTAGTCTCAGTTGTGGTCTCCA	360
DB	319	ATTTCAACGGTTGATGGATGGCTTGAGAAGTACTGCGCTAGTCTCAGTTGTGGTCTCCA	378
QY	361	TCATCATGCAATGGTGGAAAGGGTCATTTTGGGGCCACGAGTGGGAGAAACATGGACT	420
DB	379	TCATCATGCAATGGTGGAAAGGGTCATTTTGGGGCCACGAGTGGGAGAAACATGGACT	438
QY	421	TGTTCTTCTCTGTTTTCATGATGAGTATAAATTAATCTCTTACCACACTTAATCTCTAC	480
DB	439	TGTTCTTCTCTGTTTTCATGATGAGTATAAATTAATCTCTTACCACACTTAATCTCTAC	498
QY	481	TTGAAGCATAAATGTCACGGATGTCCTTTATCAAGCTGGCTATGTTGCTTCCAACAGTGA	540
DB	499	TTGAAGCATAAATGTCACGGATGTCCTTTATCAAGCTGGCTATGTTGCTTCCAACAGTGA	558
QY	541	AAGTATCCTCTAGAGAGTATCGTAAACAGCCATTACGAATGCATTTTCATATCACCCCTGAA	600
DB	559	AAGTATCCTCTAGAGAGTATCGTAAACAGCCATTACGAATGCATTTTCATATCACCCCTGAA	618
QY	601	GTGTTTGCAGAAAGAGATGCAATCGATGAAATACGTATATGCTTCTATAAAGATTTTAAG	660
DB	619	GTGTTTGCAGAAAGAGATGCAATCGATGAAATACGTATATGCTTCTATAAAGATTTTAAG	678
QY	661	CCAGGAGCTGTGTTGTTTACAAAGATTTGACATCTAGAAAGTCATGCCCAAGTAGTA	720
DB	679	CCAGGAGCTGTGTTGTTTACAAAGATTTGACATCTAGAAAGTCATGCCCAAGTAGTA	738
QY	721	AGTTTGCAGAAATACAGCCATTAGATGTTGAAGCTATGTTCTGAAGATGCCAACGAA	780
DB	739	AGTTTGCAGAAATACAGCCATTAGATGTTGAAGCTATGTTCTGAAGATGCCAACGAA	797
QY	781	AGAGAAGCTCTTTCAATCGGAAAGATGGGAGGTTTGTATCTCTCTGAGAGACATACAT	840
DB	798	AGAGAAGCTCTTTCAATCGGAAAGATGGGAGGTTTGTATCTCTCTGAGAGACATACAT	857
QY	841	ACATGTCTCTGATGTTGTAACTTTACTACAAAAACCTATAAAGATTTGGCTTATTCGTTTC	900
DB	858	ACATGTCTCTGATGTTGTAACTTTACTACAAAAACCTATAAAGATTTGGCTTATTCGTTTC	917
QY	901	TATTGG--ATATGTATCATCTACCTGTTAAATCAAGATTCTTTCTTAATATGTAGAGA	958
DB	918	TATTGGATATATGATCATCATTAATCTGGTAAATCAAGTTTCTTTCTTAATATGTAAAGA	977
QY	959	TCAGAAAAATCCATGAAGAGATATCAACATTTGAGTTCTATGTGTAATAAAAA	1009
DB	978	TCAGAAAAATCCATGAAGAGATATCAACATTTGAGTTCTATGTGTAATAAAAA	1028

RESULT 3

ID	AAQ71177	standard; cDNA; 822 BP.
AC	AAQ71177;	
XX		
XX	25-MAR-2003 (updated)	
DT	27-MAR-1995 (first entry)	
XX		
XX		
DE	E. pinatifida ribonuclease P8 cDNA.	
XX		
KW	Ribonuclease; RNAase; protein P8; pathogenic fungus;	
KW	namatode damage; crop improvement; transgenic plant;	
KW	disease-resistance; ds.	
XX		
OS	Engelmannia pinatifida.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	10..801
FT		/*tag= a
FT	mat_peptide	10..798
FT		/*tag= b
XX		
PN	W09418335-A2.	
XX		
FD	18-AUG-1994.	
XX		
XX	24-JAN-1994;	94WO-US0844.
PF		
XX		
PR	29-JAN-1993;	93US-0010403.
PR	16-DEC-1993;	93US-0166172.
XX		
XX		
PA	(MONS) MONSANTO CO.	
PI	Hironaka CM, Huynh QK, Shah DM;	
XX		
DR	WPI; 1994-279756/34.	
XX	P-PSDB; AARG0474.	
XX		
PT	Control of plant pathogenic fungi using ribonucleases - also	
PT	useful for control of nematode damage to plants	
XX		
PS	Disclosure; Page 28-29; 35pp; English.	
XX		
CC	E. pinatifida leaf cDNA expressing ribonuclease P8 was	
CC	expressed in transgenic plants to improve resistance to fungal	
CC	pathogens and to nematode attack.	
CC	(Updated on 25-MAR-2003 to correct PN field.)	
XX		
XX		
SQ	Sequence 822 BP; 228 A; 175 C; 173 G; 246 T; 0 other;	
	Query Match	16.9%; Score 170.8; DB 15; Length 822;
	Best Local Similarity	56.0%; Pred. No. 1.8e-37;
	Matches	366; Conservative 0; Mismatches 282; Indels 6; Gaps 2;
Qy	29	TATGTCCTTCCTTCGCTGGTGTATGCCGAGCATTGCCGAGACGTCATCGAAC 88
Db	41	TATCAGTTCTGCTTACCGGATTGTTCTCCGTTGATGGAGACGATGATCTCGTTCAA 100
Qy	89	TCAATCGATCTCAGAGGAGTTCGATTATTTCCTCTATCTCTTCAATGCCCTGGAACCT 148
Db	101	CTCTCCCGAACACACAAAGATTGATTCTTCACACTGGCTCTGCAATGCCAGCACTT 160
Qy	149	ATTGGCGTGGAACT---CGCCATTGTGCTCCAAAACCTTGCTCAGAGGCTCCGATG 205
Db	161	TCGTCTCAACGATGAGAACAAATGCTGTCTCTGAAATGGTTGTTGCCAAGGAGGAAT 220
Qy	206	CTCCAACTCAATTCAATTCATGGGTTATGGCCTCACTATACGATGGTTCGTGGCCTT 265
Db	221	CTCCACAGGATTTACATCCACGGACTATGGCCCATACAGTGATGGAACATGGCCAT 280

PT peptide, for manufacturing complex carbohydrates, or as targets for
PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant
PT rejection
XX
XX Disclosure; SEQ ID NO 5737; 6pp; English.
XX
XX The invention relates to a composition comprising a human GDP-mannose
CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
CC in a mammalian subject and for treating or ameliorating diseases affected
CC by the level of cellular fucosylation or diseases affected by the
CC fucosylation of glycoconjugates. These diseases include arthritis,
CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or
CC infection. The GM4,6D peptide or a polynucleotide encoding it is also
CC useful for manufacturing complex carbohydrates and as targets for
CC screening small molecule antagonists of the activity of the enzyme. The
CC polynucleotide is useful in developing an assay for defects in the
CC enzyme, as well as in gene replacement therapy. Sequences
CC ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding
CC human GM4,6D peptides of the invention. The GM4,6D peptide did not form part of the printed
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 229 BP; 49 A; 69 C; 51 G; 60 T; 0 other;
SQ
Query Match 7.6%; Score 76.6; DB 25; Length 229;
Best Local Similarity 67.5%; Pred. No. 2.3e-11; Indels 2; Gaps 2;
Matches 137; Conservative 0; Mismatches 64;
QY 99 TCAGAGGAGTTCGATTATTCGCTCTATCTTCAATGCGCTGGAACTTA-TTGGCGTG 157
DB 27 TCAACGAGAGTTGACTATATCGCTTTCGGGTTCGATGCGCGGTACTTACTCGCAAC 86
QY 158 GAATCGGCATTTGTCTCCAAACCGTTCTGCAGAGCTCGATGCTCCAACTCAT 217
DB 87 GCACCCCGAGTTGTGACCCACCGTTGCTGCAGAGCTCGATGCTCCAACTCAT 146
QY 218 TCACAAATTCATGGTTATGGCTGACTATACGATGCTGGCGCC-TTCATGTTGTAT 276
DB 147 TCACATACATGACTCTGGCTGACTATATGATAGACCTGGCGCTACCTGTGTCT 206
QY 277 CGATCTGACTTTAAAGAGAGGA 299
DB 207 GGAATCTAGTTTCGATCTCTAAAGA 229
RESULT 6
ABX26579
ID ABX26579 standard; cDNA; 224 BP.
XX
XX AC ABX26579;
XX
XX 11-FEB-2003 (first entry)
XX
XX Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #8636.
XX
XX Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
XX cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
XX arthritis; asthma; sepsis; reperfusion injury; stroke; infection;
XX complex carbohydrate; gene replacement therapy; immunosuppressive;
XX antiinflammatory; antiarthritic; antibacterial; cerebroprotective;
XX antiasthmatic; vasotropic.
XX
XX Homo sapiens.
XX
XX US2002110548-A1.
XX
XX 15-AUG-2002.
XX
XX 11-JUN-2001; 2001US-0878574.
XX
XX 22-NOV-1996; 96US-0753233.
XX

PR 03-DEC-1997; 97US-0984246.
PR 09-SEP-1998; 98US-0149674.
PR 14-JUN-1999; 99US-0333177.
XX
XX (GEMY) GENETICS INST INC.
XX
XX Sullivan F, Kriz R, Kumar R;
XX WPI; 2003-066673/06.
XX
XX New composition comprising GDP-mannose 4,6-dehydratase (GM4,6D)
PT peptide, for manufacturing complex carbohydrates, or as targets for
PT screening GM4,6D antagonists for treating e.g. arthritis, or transplant
PT rejection
XX
XX Disclosure; SEQ ID NO 8638; 6pp; English.
XX
XX The invention relates to a composition comprising a human GDP-mannose
CC 4,6-dehydratase (GM4,6D) peptide. The peptide is useful for identifying
CC GM4,6D inhibitors. GM4,6D inhibitors are useful for reducing inflammation
CC in a mammalian subject and for treating or ameliorating diseases affected
CC by the level of cellular fucosylation or diseases affected by the
CC fucosylation of glycoconjugates. These diseases include arthritis,
CC transplant rejection, asthma, sepsis, reperfusion injury, stroke or
CC infection. The GM4,6D peptide or a polynucleotide encoding it is also
CC useful for manufacturing complex carbohydrates and as targets for
CC screening small molecule antagonists of the activity of the enzyme. The
CC polynucleotide is useful in developing an assay for defects in the
CC enzyme, as well as in gene replacement therapy. Sequences
CC ABX17942-ABX17944 and ABX17947-ABX33716 represent DNA molecules encoding
CC human GM4,6D peptides of the invention. The GM4,6D peptide did not form part of the printed
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 224 BP; 37 A; 71 C; 53 G; 63 T; 0 other;
SQ
Query Match 6.3%; Score 63.8; DB 25; Length 224;
Best Local Similarity 69.9%; Pred. No. 8.5e-06;
Matches 86; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 99 TCAGAGGAGTTCGATTATTCGCTCTATCTTCAATGCGCTGGAACTTA-TTGGCGTG 158
DB 102 TCAACGAGAGTTGACTATTCGCTTTCGCTTTCGATGCGCGGTACTTACTCGCAACG 161
QY 159 AACTCGCATTTGTCTCCAAACCGTTGCTGCAGAGCTCGATGCTCCAACTCAT 218
DB 162 CACCGCGAGTTTCTGCCCCATCATCGTTGCTGCAGAGCTCGATGCTCCAACTCAT 221
QY 219 CAC 221
DB 222 CAC 224
RESULT 7
AAC34348
ID AAC34348 standard; DNA; 1030 BP.
XX
XX AAC34348;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 6327.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EF1033405-A2.
XX
XX 06-SEP-2000.
XX


```
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162143.

Query Match 5.0%; Score 50.4; DB 21; Length 1030;
Best Local Similarity 52.5%; Pred. No. 0.00091;
Matches 214; Conservative 0; Mismatches 176; Indels 18; Gaps 4;

QY 212 CTCATTCACAAATTCATGGTGTATGGCTGACTATACGATGGTTCGTGGCCCTTCATGTT 271
DB 387 CTGATTTGGCATTCATGGTCTTTGGCTTAACACAAAGATGGCACTTATCCATCTAACT 446
QY 272 GTTATCGATCTGATTTAAAGAGAGAGATTTCAACGTTGATGGATGGTCTTGAGAGT 331
DB 447 GTGA-----TGACTCTAAACCTTTCGATAGCTCAACGATATCAGATCTTTCT-----CA 494
QY 332 ACTGCGCTAGTCTCAGTTGTGGTTCCTCATCATCATGCAATGGTGGGAAAGGTCATTTT 391
DB 495 CCTCGATGAAGAGAGCTGGCCAAACACTGGCTTCCCAAGCGTTCAGGTGAAGCGTTT 554
QY 392 GGGGCCAGAGTGGGAGAAAATGGGACTTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 451
DB 555 GGGAGCAGCAATGGGAGAGCATGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 611
QY 452 ATTACTCTCTTACCACCTTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 511
DB 612 AATATTTCCAAACCGCTTTAACTTAAAGAGAAACCAATCTCTCTCTCTCTCTCTCTCTCT 671
QY 512 AAGCTGGCTATGTTGCTTCCAAAGTGAAGATGATGATGATGATGATGATGATGATGATGAT 571
DB 672 AAGCCGGGA---TTAATCCGGATGGAATCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 728
QY 728 TTCAGATGCAATTCATATCACTCCCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 619
DB 729 TAAAGAGTCAATTCATGTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 776

ABZ14066 standard; DNA; 693 BP.
ABZ14066;
21-JAN-2003 (first entry)
Arabidopsis thaliana stress regulated gene SEQ ID NO 1871.
Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX WO2000216655-A2.
XX 28-FEB-2002.
XX 24-AUG-2001; 2001WO-US26685.
XX 24-AUG-2000; 2000US-227866P.
XX 26-JAN-2001; 2001US-264647P.
XX 22-JUN-2001; 2001US-300111P.
XX (SCRI ) SCRIPPS RES INST.
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX Harper JF, Kreps J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
XX Identifying a stress condition to which a plant cell has been exposed
XX and producing plants with increased tolerance to these abiotic stresses
XX
XX Claim 144; SEQ ID NO 1871; 577pp + Sequence Listing; English.
XX The invention relates to identifying a stress condition to which a plant
XX cell has been exposed, comprising:
XX (a) contacting nucleic acid representative of expressed polynucleotides
XX in the plant cell with an array or probes representative of the plant
XX cell genome; and
XX (b) detecting a profile of expressed polynucleotides in the plant cell
XX characteristic of a stress response. The method is useful in the
XX production of transgenic plants, cells and seeds and in producing plants
XX with increased tolerance to abiotic stress. The present sequence is that
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
XX in methods of the invention.
XX Note: The sequence data for this patent is not represented in the printed
XX specification but is based on sequence information supplied to Derwent by
XX the European Patent Office.
XX Sequence 693 BP; 182 A; 150 C; 145 G; 216 T; 0 other;

Query Match 4.6%; Score 46.8; DB 24; Length 693;
Best Local Similarity 56.0%; Pred. No. 0.0078;
Matches 131; Conservative 0; Mismatches 97; Indels 6; Gaps 2;

QY 386 CATTTTGGGCGCAGAGTGGGAGAACATGGGACTTGTCTCTCTCTCTCTCTCTCTCTCTCT 445
DB 341 CGTTTGGGAGCAGCAATGGGAGAGCATGGTACTTGTCTCTCTCTCTCTCTCTCTCTCTCT 397
QY 446 AGTATAATTACTTCTCTTACCACATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 505
DB 398 AACATGAATATTCCAAACCGCTTTAACTTAAACAGAAACCAATCTCTCTCTCTCTCTCT 457
QY 506 TTTATCAAGCTGGGTATGTTGTTCCAAAGTGAAGATCTCTCTCTCTCTCTCTCTCTCTCT 565
DB 458 TAAACAAAGCCGGGA---TTAATCCGGATGGAATCTTACTCTCTCTCTCTCTCTCTCTCT 514
QY 566 CAGCCATTGAGATGATTTTATATACCCCTGGAAGTGGTGGTGGTGGTGGTGGTGGTGGT 619
DB 515 ATTCTGATAAAGAGTCAATTTGTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 568

RESULT 8
ABZ14066
```


RESULT 9
AAC47291
ID AAC47291 standard; DNA; 1026 BP.
XX AC AAC47291;
XX DT 18-OCT-2000 (first entry)
XX XX Arabidopsis thaliana DNA fragment SEQ ID NO: 53278.
DE XX Arabidopsis thaliana
XX XX Hybridisation assay; genetic mapping; gene expression control;
KW KW protein identification; signal transduction pathway;
KW KW metabolic pathway; promoter; termination sequence; ss.
XX XX Arabidopsis thaliana.
XX XX EP1033405-A2.
XX XX 06-SEP-2000.
XX XX 25-FEB-2000; 2000EP-0301439.
XX XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	26-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151203.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0158293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	23-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 4.6%; Score 46.8; DB 21; Length 1026;
 Best Local Similarity 56.0%; Pred. No. 0.0092;
 Matches 131; Conservative 0; Mismatches 97; Indels 6; Gaps 2

QY	386	CATTTTGGGCGCACAGTGGGAGAAAATGGCACTTCCTCTCCTGTGTTTCATGATG	445
Db	549	CGTTTTGGAGACGAATGGGAGAGCATGTACTTCTCTGTAATCGGTAT---CGATC	605
QY	446	AGTAATAATTACTCTCTTACCACACTTAATCTCTACTTGAAGCATATGTCACGGATGTC	505
Db	606	AACAAAGCCGGG---TTAATCGGATGAAAACTTACTCTTTTGAGAGCATAAGAG	722
QY	506	TTTATCAAGCTGGCTATGTGTCTTCAAACAGTGAAGAAGTATCCTCTAGGAGGTATCGTAA	565
Db	666	TAAACAAAGCCGGG---TTAATCGGATGAAAACTTACTCTTTTGAGAGCATAAGAG	722

Qy	566	CAGCCATTGAGAAATGCAATTCATATCAACCCCTGAAGTGGTTTCAAAAGAGATG	619
Db	723	ATTTCGATAAAGAGTCAATGGTTTCACTCCTTGGGTTGAGTGAACAGAGATG	776
RESULT 10			
ABN71527/c			
ID	ABN71527	standard; DNA; 2155561 BP.	
XX	AC	ABN71527;	
XX	XX		
XX	DT	02-JUL-2002 (first entry)	
XX	XX		
XX	DE	Streptococcus polynucleotide SEQ ID NO 10967.	
XX	XX		
XX	KW	Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;	
KW	group A streptococcus; Streptococcus pyogenes; antibacterial; gene;		
KW	anti-inflammatory; infection; vaccine; meningitis; gene therapy; ds.		
XX	OS	Streptococcus sp.	
XX	XX		
XX	FN	WO200234771-A2.	
XX	XX		
XX	PD	02-MAY-2002.	
XX	XX		
XX	PF	29-OCT-2001; 2001WO-GB04789.	
XX	XX		
XX	PR	27-OCT-2000; 2000GB-0026333.	
PR	24-NOV-2000; 2000GB-0028727.		
PR	07-MAR-2001; 2001GB-0005640.		
XX	XX		
XX	PA	(CHIR-) CHIRON SPA.	
PA	PA	(GENO-) INST GENOMIC RES.	
XX	XX		
PI	Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;		
PI	Tettelin H;		
XX	XX		
DR	WPI: 2002-352536/38.		
XX	XX		
PT	New Streptococcus protein for the treatment or prevention of infection		
PT	or disease caused by Streptococcus bacteria, such as meningitis, and		
PT	for detecting a compound that binds to the protein -		
XX	XX		
PS	Claim 8; Page 4196-4488; 4525pp; English.		
XX	XX		
CC	The invention relates to a protein (AB25413-ABP30895) from group B		
CC	streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS		
CC	(Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in		
CC	the specification. The proteins have antibacterial and anti-inflammatory		
CC	activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and		
CC	antibodies that bind (I) are used in the manufacture of medicaments for		
CC	the treatment or prevention of infection or disease caused by		
CC	Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.		
CC	Nucleic acids encoding (I) are used to detect Streptococcus in a		
CC	biological sample. (I) is used to determine whether a compound binds to		
CC	(1). A composition comprising (I) or a nucleic acid encoding (I), may be		
CC	used as a vaccine or diagnostic composition. The disease caused by		
CC	Streptococcus that is prevented or treated may be meningitis. Nucleic		
CC	acid encoding (I) may be used to recombinantly produce (I) and may be		
CC	used in gene therapy. Antibodies to (I) are used for affinity		
CC	chromatography, immunoassays, and distinguishing/identifying		
CC	Streptococcus proteins.		
XX	XX		
SQ	Sequence 2155561 BP; 695741 A; 385320 C; 382958 G; 691542 T; 0 other;		
Query Match	4.3%; Score 43.6; DB 24; Length 2155561;		
Best Local Similarity	55.2%; Pred. No. 1.9;		
Matches	85; Conservative 0; Mismatches 69; Indels 0; Gaps 0;		
Qy	856	TGTAACCTTTACTACCAAAACCTATAAAGATGGCTTATTCGTTCTATTGGATATGATC	915
Db	1665003	TATAAGTCGAATCTAGAAATTCCAATTTTCATTTTATTTTCTATTGTTATTTATTC	1665

QY 916 ATCATTAACGTAAGTCTTCTTCTTAATAGTAGAAGATCAGAAATCCATAGA 975
 Db 1664943 TTTATTGATCTAGTTTTCAGGATTTACAGTAAATGATGAGAGAGAAATGATTATGA 1664884
 QY 976 AGATATCAACATTTGAGTCTTCTTCTTAATAGTAGAAGATCAGAAATCCATAGA 1009
 Db 1664883 AGAGCTAGAAATTTTCTCCTGACAGATTAAGAAAA 1664850

RESULT 11
 ID ABK31510/C
 ABK31510 standard; DNA; 47108 BP.

XX ABK31510;

XX 23-APR-2002 (first entry)

XX Signal transduction associated gene modified DNA #177.

XX Human; signal transduction associated gene; cytosine methylation state;
 KW CpG island; signal transduction associated disease; solid tumour; cancer;
 KW antitumour; cytostatic; mutant; ds.

XX Homo sapiens.

OS Synthetic.

XX WO200200926-A2.

XX 03-JAN-2002.

XX 29-JUN-2001; 2001WO-EP07472.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-147896/19.

XX Oligonucleotide for diagnosis and therapy of diseases associated with
 PT signal transduction e.g. cancer, comprises chemically modified genomic
 PT sequences of genes associated with signal transduction -

PS Claim 1; SEQ ID No 353; 24pp; English.

CC The present invention relates to chemically modified DNA sequences of
 CC signal transduction associated genes. The DNA sequences are chemically
 CC modified using a solution of bisulphite, hydrogen sulphite or
 CC disulphite. Also disclosed are oligonucleotides and/or RNA oligomers
 CC for detecting the cytosine methylation state (CpG islands) of these
 CC genes, and a method for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with signal transduction.
 CC The genomic DNA can be obtained from cells or cellular components which
 CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
 CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
 CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
 CC histologic object slides, and all their possible combinations. The
 CC sequences of the invention are useful for the diagnosis and therapy of
 CC diseases associated with signal transduction e.g. solid tumours and
 CC cancer. ABK3158-ABK31545 represent chemically pretreated genomic DNA
 CC sequences of different genes associated with signal transduction, or
 CC their complementary sequences.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

XX Sequence 47108 BP; 13283 A; 429 C; 9740 G; 23656 T; 0 other;

Query Match 4.2%; Score 42.6; DB 24; Length 47108;

Best Local Similarity 55.8%; Pred. NO. 0.7;

Matches 101; Conservative 0; Mismatches 79; Indels 1; Gaps 1;

QY 833 CAATACATACATGCTCTGATGTTGTAACTTATACCAAAACCTATAAAGATTGGCTTA 892
 Db 7217 CAATACCTTACATAAACATTTATTTCTTTTAACTTCCCCCAAAAACCAAAATA 7158
 QY 893 TTTCGTTCTATTGGATATGATCATCATAC-TGGTAAATCAAGTTTCTTTCTTAATATG 951
 Db 7157 ACATTACCATAACATATATATAATAAACAATATAAATTAATATCAATAACC 7098
 QY 952 TAGAAGATCAGAAATCCATAAGAGATATCAACATTGAGTTCTTATGTTAAAAAAA 1011
 Db 7097 TAAAAAATCAGCAAAACCAAAATAATATTTATTTATTTATTTATTTAAAAACAA 7036
 QY 1012 A 1012
 Db 7037 A 7037

RESULT 12

ABL49339

ID ABL49339 standard; DNA; 19513 BP.

XX ABL49339;

XX 01-MAY-2002 (first entry)

XX Human polynucleotide associated with DNA replication SEQ ID NO 39.

XX Human; cytostatic; neuroprotective; nootropic; immunostimulant;

KW gene therapy; gene regulation; DNA replication; CENPB; DNA2L; ATR; CHD1L;

KW ERCC3; SNRPA1; RAD50; LIG2; cytosine methylation; Ataxia telangiectasia;

KW ATR-X; Bloom's syndrome; tumour; cancer; methylation; gene; ds.

XX Homo sapiens.

XX WO200177377-A2.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-EP03971.

XX 06-APR-2000; 2000DE-1019058.

XX 07-APR-2000; 2000DE-1019173.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-017471/02.

XX New nucleic acid sequences from chemically modified genes associated
 PT with DNA replication, useful for analysing cytosine methylations for
 PT diagnosis and therapy of diseases e.g. Ataxia telangiectasia -

XX Claim 1; SEQ ID NO 39; 23pp + Sequence Listing; English.

XX The invention relates to nucleic acid sequences comprising at least 18
 CC bases of a chemically pretreated gene associated with gene regulation,
 CC selected from 94 genes (ABL49301-ABL49394) and/or complementary sequences
 CC associated with DNA replication, CENPB, DNA2L, ATR, CHD1L, ERCC3, SNRPA1,
 CC RAD50 and LIG2. The chemical pretreatment converts cytosine bases
 CC unmodified at the 5-position to uracil or another base with
 CC hybridisation behaviour dissimilar to cytosine, to enable analysis of
 CC cytosine methylations. The DNA sequences and method are useful in the
 CC diagnosis of diseases (or predisposition to diseases) associated with DNA
 CC replication and in therapy of such diseases. They are especially useful
 CC cytosine methylation patterns of such genes. They are especially useful
 CC in diagnosis and therapy of e.g. Ataxia telangiectasia, ATR-X, Bloom's
 CC syndrome, solid tumours and cancer.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification and was supposed to be available directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences. However, the sequence data did
 CC not correspond to that referred to in the specification. The present data
 CC is taken from BPO data for the patent.
 XX
 SQ Sequence 19513 BP; 4543 A; 563 C; 5296 G; 9111 T; 0 other;
 Query Match 4.0%; Score 40.8; DB 24; Length 19513;
 Best Local Similarity 50.5%; Pred. No. 1.5;
 Matches 99; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
 QY 789 TCTTTGAATCGGAAAGATGGAGCTTTGTATCTCTCTGAGAGACATACATACATGCT 848
 Db 9107 TTTTGTATTATAATAAAGACGCTTTATATATTTTCTTATTATAATAAAGACGTT 9166
 QY 849 CTGATGTTGTAACCTTACTACCAAAACCTATAAGATGGCTTATCTCTCTTATGGAT 908
 Db 9167 TATATATTTTGTATTAAGTAAAGATTTTATATATTTTGTATTATAATAAAGACGTT 9226
 QY 909 ATGTATCATCATCTACTGTAATCAAGTTCTTTCTTAATAATGATAGAGATCAGAAATC 968
 Db 9227 ATTGTATTATTTTATTATATATTTTATATATTTTATATATTTTGAAGATTTTAAATTT 9286
 QY 969 CATAGAAGATATCAA 984
 Db 9287 GTAGAGAAGGTATAAA 9302
 RESULT 13
 AAS61245
 ID AAS61245 standard; DNA; 19513 BP.
 AC AAS61245;
 XX
 XX
 DT 29-JAN-2002 (first entry)
 XX
 DE Human gene regulation-associated gene oligonucleotide #200.
 KW Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preclampsia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
 KW immunostimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
 KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cytostatic.
 KW
 OS Homo sapiens.
 XX
 XX WO200177375-A2.
 XX
 XX 18-OCT-2001.
 PD
 XX
 XX 06-APR-2001; 2001WO-EP03968.
 XX
 XX 06-APR-2000; 2000DE-1019058.
 PR
 XX 07-APR-2000; 2000DE-1019173.
 PR
 XX 30-JUN-2000; 2000DE-1032529.
 PR
 XX 01-SEP-2000; 2000DE-1043826.
 XX
 XX (EPIG-) EPIGENOMICS AG.
 PA
 XX Olek A, Piepenbrock C, Berlin K;
 PI
 XX WPI; 2002-017470/02.
 DR
 XX
 XX New nucleic acid sequences from chemically modified genes associated
 PT with gene regulation, useful for analysing cytosine methylations for
 PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
 PT disease
 PT
 XX Claim 1; SEQ ID No 205; 26pp; English.
 PS
 XX The invention relates to 224 nucleic acid sequences comprising at least
 XX 18 bases of a chemically pretreated gene associated with gene regulation
 CC

CC selected from 43 known genes (or complementary sequences). The
 CC chemical pretreatment converts cytosine bases unmethylated at the
 CC 5-position to uracil or another base with hybridisation behaviour
 CC dissimilar to cytosine, to enable analysis of cytosine methylations.
 CC The DNA sequences, oligomers (or sets/arrays) and method are
 CC useful in the diagnosis of diseases (or predisposition to diseases),
 CC associated with gene regulation and in therapy of such diseases, by
 CC enabling analysis of the cytosine methylation patterns of such genes,
 CC kits are provided. They are especially useful in diagnosis
 CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
 CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
 CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
 CC preclampsia, graft versus-host disease. The present sequence is a
 CC sequence included in the sequence data for this specification and is
 CC associated with the human gene regulation-associated genes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 19513 BP; 4543 A; 563 C; 5296 G; 9111 T; 0 other;
 Query Match 4.0%; Score 40.8; DB 24; Length 19513;
 Best Local Similarity 50.5%; Pred. No. 1.5;
 Matches 99; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
 QY 789 TCTTTGAATCGGAAAGATGGAGCTTTGTATCTCTCTGAGAGACATACATACATGCT 848
 Db 9107 TTTTGTATTATAATAAAGACGCTTTATATATTTTGTATTATAATAAAGACGTT 9166
 QY 849 CTGATGTTGTAACCTTACTACCAAAACCTATAAGATGGCTTATCTCTCTTATGGAT 908
 Db 9167 TATATATTTTGTATTAAGTAAAGATTTTATATATTTTGTATTATAATAAAGACGTT 9226
 QY 909 ATGTATCATCATCTACTGTAATCAAGTTCTTTCTTAATAATGATAGAGATCAGAAATC 968
 Db 9227 ATTGTATTATTTTATTATATATTTTATATATTTTATATATTTTGAAGATTTTAAATTT 9286
 QY 969 CATAGAAGATATCAA 984
 Db 9287 GTAGAGAAGGTATAAA 9302
 RESULT 14
 ABL32458/c
 ID ABL32458 standard; DNA; 10133 BP.
 XX
 XX ABL32458;
 AC
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX
 XX Human immune system associated gene SEQ ID NO: 431.
 DE
 XX Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antianaemic; cytostatic; neurotropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 KW
 OS Homo sapiens.
 XX
 XX WO200200928-A2.
 PN
 XX
 XX 03-JAN-2002.
 PD
 XX
 XX 02-JUL-2001; 2001WO-EP07537.
 PF
 XX
 XX 30-JUN-2000; 2000DE-1032529.
 PR
 XX 01-SEP-2000; 2000DE-1043826.
 PR
 XX

PA (BPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-130909/17.
 XX
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation -
 XX
 PS Claim 1; SEQ ID NO 431; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 10133 BP; 3075 A; 70 C; 2051 G; 4937 T; 0 other;
 Query Match 4.0%; Score 40.4; DB 24; Length 10133;
 Best Local Similarity 52.4%; Pred. No. 1.5;
 Matches 89; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
 QY 843 ATGCTCTGATGTTGTAACCTTTACTACCAAAACCTATAAAGATTGCTTATTTCGTTCTA 902
 DB 2924 ATAACAATTACGTACAAATATATAATTACAAATATATATATATATATATATATATAA 2865
 QY 903 TTGGATATGATCATCATATTACTGTTAAATCAAGTTTCTTTCTTAATAATGTAAGATCAG 962
 DB 2864 CTACAAATATATATACATATTATATAAAACACATACCTAACCTAAACCAAAATTCAAA 2805
 QY 963 AAAATCATAGAGATATCAACATTGATTTCTTCTGTTAAACCAAA 1012
 DB 2804 ACAACCTAAACACATACAAACCTTATCTCTACTACCAAAACCAAA 2755
 RESULT 15
 ABL55643/C
 ID ABL55643 standard; DNA; 50000 BP.
 XX
 AC ABL55643;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE AmEPV genome fragment#1.
 XX
 KW AmEPV; gene therapy; viral vector; chromosome mapping; gene mapping;
 KW Genetic deficiency disorder; ds.
 XX
 OS Amsacta moorei entomopoxvirus.
 XX
 PN WO200212526-A2.
 XX
 PD 14-FEB-2002.
 XX
 PF 10-AUG-2001; 2001WO-US25287.
 XX
 PR 10-AUG-2000; 2000US-224479P.
 PR 14-SEP-2000; 2000US-0662254.
 XX
 XX (UYFL) UNIV FLORIDA.
 XX
 PI Moyer²RW, Li Y, Bawden AL;
 XX
 DR WPI; 2002-227161/28.
 XX
 XX Novel recombinant entomopox virus vector useful for delivering
 PT polynucleotide encoding protein to vertebrate cell, comprises
 PT polynucleotide encoding protein operably linked with heterologous

PT promoter sequence -
 XX
 PS Disclosure; Page 125-150; 326pp; English.
 XX
 CC The invention relates to a recombinant entomopox virus (EPV) vector,
 CC comprising a polynucleotide encoding a protein operably linked with a
 CC heterologous promoter sequence. The invention also concerns methods for
 CC providing gene therapy for genetic deficiency disorders. Vectors of the
 CC invention are useful for delivering a polynucleotide encoding a protein
 CC to a vertebrate cell preferably a mammalian cell, such as a human cell.
 CC The vector is introduced into the vertebrate cell by infection in a viral
 CC particle, or by transfection, transduction, or injection either in vitro
 CC or in vivo. The vector is useful for the delivery and expression of
 CC biologically useful proteins in gene therapy protocols, and for
 CC delivering large DNA segments for engineering of vertebrate cells.
 CC Polynucleotides of the invention have applications in techniques such as
 CC their use as insertion sites for foreign genes of interest, hybridisation
 CC probes, for chromosome and gene mapping, in PCR technologies, and in the
 CC production of sense or antisense nucleic acids. Vectors of the invention
 CC provide for stable integration and expression of heterologous DNA in host
 CC cells, and are adapted for accepting large heterologous polynucleotide
 CC inserts which can be delivered in an infected or transformed cell and
 CC expressed in a stable fraction. The current sequence represents a
 CC fragment of the genome of the genus B entomopoxvirus from amsacta moorei
 CC (AmEPV).
 XX
 SQ Sequence 50000 BP; 19857 A; 4373 C; 4222 G; 21548 T; 0 other;
 Query Match 4.0%; Score 40.2; DB 24; Length 50000;
 Best Local Similarity 51.4%; Pred. No. 3.4;
 Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
 QY 823 TTCTGAGAGACAATACATACATGCTCTGATGTTGTAACTTTACTACCAAAACCTATAAA 882
 DB 30790 TTATTATAGGCTATATAAATTATTTAAATAATTATCGTTACAATTATAATTATAT 30731
 QY 883 GATTGGCTTATTCGTTCTTATGATATGATCATCATCTACTGTAATCAAGTTTCTTT 942
 DB 30730 AATAATAATTTTTTTTGTATATAAATTTTATTTATACATATTTTATATGATATATCT 30671
 QY 943 CTAATAATGTAGAAGATCAGAAAATCCATAAGAAGATATCAACATTTGAGTTCTATGTA 1002
 DB 30670 GTATTTTTTTTATATAAATCTGAATATATTGACTAGTAACATGATTATTATTATAATT 30611
 QY 1003 A 1003
 DB 30610 A 30610
 Search completed: January 21, 2004, 17:31:58
 Job time : 267.957 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 15:50:10 ; Search time 340.379 Seconds
(without alignments)
10468.497 Million cell updates/sec

Title: US-09-936-869-1

Perfect score: 1320

Sequence: 1 aaactgacagcaaatatga.....tggtgctgtgcccattg 1320

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_19Jun03.*
1: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1980.DAT.*
2: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1981.DAT.*
3: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1982.DAT.*
4: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1983.DAT.*
5: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1984.DAT.*
6: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1985.DAT.*
7: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1986.DAT.*
8: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1987.DAT.*
9: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1988.DAT.*
10: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1989.DAT.*
11: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1990.DAT.*
12: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1991.DAT.*
13: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1992.DAT.*
14: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1993.DAT.*
15: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1994.DAT.*
16: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1995.DAT.*
17: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1996.DAT.*
18: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1997.DAT.*
19: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1998.DAT.*
20: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA1999.DAT.*
21: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA2000.DAT.*
22: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA2001A.DAT.*
23: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA2001B.DAT.*
24: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA2002.DAT.*
25: /SIDSL/gcgdata/Geneseq/Geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1320	100.0	1320	21 AAA99362	Plant PrAgI promot
2	4320	100.0	1401	21 AAA99363	Plant PrAgI promot
C 3	46.6	3.5	540	24 ABQ18618	Oligonucleotide fo
4	46.6	3.5	540	24 ABQ18619	Oligonucleotide fo
C 5	46.4	3.5	4492	23 ABL15002	Drosophila melanog
C 6	45.6	3.5	563	23 ABV50724	Human prostate exp
C 7	45	3.4	519	25 ABX22233	Human GDP-mannose
8	44.4	3.4	15923	24 ABL70179	Chemically treated

9	44.4	3.4	15923	24 AAS61132	Human gene regulat
10	44.4	3.4	15923	24 AKK31220	Signal transductio
11	44	3.3	6261	22 AAS46643	Tumour suppressor
12	44	3.3	6261	24 AAS61046	Human gene regulat
13	44	3.3	6261	24 ABK31160	Signal transductio
C 14	43.6	3.3	11084	20 ABK32302	Human IL-1ra BAC c
15	43.4	3.3	409	22 AAI32030	Human polynucleoti
16	43.4	3.3	13326	24 ABL33712	Human immune syste
C 17	43.2	3.3	18434	24 ABL34006	Human immune syste
C 18	43	3.3	201	22 ABA48463	Human breast cell
C 19	43	3.3	201	22 ABA51474	Human breast cell
C 20	43	3.3	201	22 ABA66368	Human foetal liver
C 21	43	3.3	201	22 ABA69502	Human foetal liver
C 22	43	3.3	201	22 ABA33430	Probe #11896 for g
C 23	43	3.3	201	22 ABA36424	Probe #14890 for g
C 24	43	3.3	201	22 AAK14785	Human brain expres
C 25	43	3.3	201	22 AAK17754	Human brain expres
C 26	43	3.3	201	22 AAK40526	Human bone marrow
C 27	43	3.3	201	22 AAK43573	Human bone marrow
C 28	43	3.3	201	22 AAI21285	Probe #11218 for g
C 29	43	3.3	201	22 AAI24360	Probe #14293 for g
C 30	43	3.3	201	22 AAI46560	Probe #15246 used
C 31	43	3.3	201	22 AAI2630	Probe #18316 used
C 32	43	3.3	201	22 AAI06989	Probe #980 used t
C 33	43	3.3	201	22 AAI09899	Probe #980 used t
C 34	43	3.3	201	23 ABS40089	Human liver single
C 35	43	3.3	201	23 ABS43218	Human liver single
C 36	43	3.3	201	24 ABS17726	Human genome-deriv
C 37	43	3.3	395	22 ABA46372	Human breast cell
C 38	43	3.3	395	22 ABA56938	Human foetal liver
C 39	43	3.3	395	22 ABA68550	Probe #5016 for ge
C 40	43	3.3	395	22 AAK05033	Human brain expres
C 41	43	3.3	395	22 AAK05063	Human bone marrow
C 42	43	3.3	395	22 AAI15180	Probe #5113 for ge
C 43	43	3.3	395	22 AAI36508	Probe #5194 used t
C 44	43	3.3	395	22 AAI04920	Probe #4911 used t
C 45	43	3.3	395	23 ABS30228	Human liver single

ALIGNMENTS

RESULT 1
AAA99362
ID AAA99362 standard; DNA; 1320 BP.
XX AC AAA99362;
XX DT 22-JAN-2001 (first entry)
XX DE Plant PrAgI promoter DNA sequence.
XX KW Plant promoter; PrAgI; reproductive tissue; transgenic plant; cereal; ds.
XX OS Pinus radiata.
XX PN WO200055172-A1.
XX PD 21-SEP-2000.
XX PF 17-MAR-2000; 2000WO-NZ00031.
XX RR 17-MAR-1999; 99NZ-0334715.
XX (CARTER HOLF-HARVEY LTD.
XX (UNMT) TASMAN BIOTECHNOLOGY LTD.
XX (UNMT) UNIV MICHIGAN TECHNOLOGICAL.
XX PI Podila GK, Liu J, Karnosky DP;
XX WPI: 2000-594442/56.
XX P-PSDB; AAB26795.

xx Novel plant reproductive tissue promoter, useful to produce plants
xx PT which have a diminished reproductive capacity or which are sterile
xx PT
xx Claim 3: Fig 2: 5100; English.
xx ps


```

PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE ) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB70899.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 39488; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57137-ABB57072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 4492 BP; 1158 A; 1245 C; 1141 G; 948 T; 0 other;
XX
XX Query Match 3.5%; Score 46.4; DB 23; Length 4492;
XX Best Local Similarity 49.6%; Pred. No. 0.068;
XX Matches 119; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
XX
QY 804 TCGACCCGACCGCTCATCTTCTTCTCTGCGCAATTTTCTGTCGATCATCATCA 863
DB 2878 TCCAGCGGATGCTATCAATCTATCGCCACCAACGATGCCATCGGATCATGTCA 2819
QY 864 TTACCATATCGGCATCCCAACCATCATCATCATCATCATCATCATCATCATCAT 923
DB 2818 TCTTCGTCTCTCATCGCGCTCAGCTCTGTCATCATCGTCGCGCGCATGTGGATGCCA 2759
QY 924 ATCGATTGTACAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 983
DB 2758 ACCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2699
QY 984 AACCAATCTGATCAGCCCTGCTCAATCTTCTTATAAATCTTATCAATCTTCAATCAT 1043
DB 2698 TAAAAAGCAACCAACCAAGCTTTAAAAATATTAGAAATATCATGCTACTGTTTATCAT 2639

RESULT 6
ABV50724/c
ID ABV50724 standard; cDNA; 563 BP.
XX
XX ABV50724;
XX
XX 17-SEP-2002 (first entry)
XX
XX Human prostate expression marker cDNA 50715.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US05171.
XX
XX 17-FEB-2000; 2000US-183319P.
XX
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
XX prostate cells and correlating with presence of prostate cancer, useful
XX for detecting presence of prostate cancer, stage of prostate cancer -
XX
XX Claim 1; Page 9866; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX specification or its complement. (I) is useful for:
XX (a) assessing whether a patient is afflicted with prostate cancer;
XX (b) monitoring the progression of prostate cancer in a patient;
XX (c) assessing the efficacy of a test compound to inhibit prostate
XX cancer in a patient;
XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX in a patient;
XX (e) selecting a composition for inhibiting prostate cancer in a patient;
XX (f) assessing the prostate cell carcinogenic potential of a compound;
XX (g) determining whether prostate cancer has metastasized in a patient;
XX (h) assessing the aggressiveness or indolence of prostate cancer in a
XX patient;
XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
XX Sequence 563 BP; 45 A; 107 C; 22 G; 385 T; 4 other;
XX
XX Query Match 3.5%; Score 45.6; DB 23; Length 563;
XX Best Local Similarity 49.2%; Pred. No. 0.047;
XX Matches 120; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
XX
QY 267 AATTATTAAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 326
DB 520 AAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 461
QY 327 GTGGGCGAGAAATATACACAGAGAGTACTTTAAACAATCGCAACCAAGGCGAGTTTCA 386
DB 460 GACAGAAAGGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 401
QY 387 ACTTGATTCTGGAGCTCGAATACGAGATATGTTGGTAAAGATAAAGGAGAGTGGAG 446
DB 400 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 341
QY 447 TGCATTTGAAATGAATGAGAGCGCACAAATGAGGAGCAATAAATGAAATATAATGC 506
DB 340 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 281
QY 507 AAGA 510
DB 280 AAAA 277

RESULT 7
ABX22233/c
ID ABX22233 standard; cDNA; 519 BP.
XX
XX ABX22233;
XX
XX 10-FEB-2003 (first entry)
XX
XX Human GDP-mannose 4,6-dehydratase (GM4,6D) DNA #4290.
XX
XX Human; GDP-mannose 4,6-dehydratase; GM4,6D; gene; ss; inflammation;
XX cellular fucosylation; glycoconjugate fucosylation; transplant rejection;
XX

```


CC genes, and a method for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with signal transduction.
 CC The genomic DNA can be obtained from cells or cellular components which
 CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
 CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
 CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
 CC histologic object slides, and all their possible combinations. The
 CC sequences of the invention are useful for the diagnosis and therapy of
 CC diseases associated with signal transduction e.g. solid tumours and
 CC cancer. ABK1159-ABK31545 represent chemically pretreated genomic DNA
 CC sequences of different genes associated with signal transduction, or
 CC their complementary sequences.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC European Patent Office.

XX SQ Sequence 15923 BP; 3153 A; 441 C; 4469 G; 7860 T; 0 other;
 Query Match 3.4%; Score 44.4; DB 24; Length 15923;
 Best Local Similarity 50.8%; Pred. No. 0.4;
 Matches 133; Conservative 0; Mismatches 126; Indels 3; Gaps 1;
 QY 97 GTACTCAGGAAGGTAGTGGACTCTCTCAGGTACATAGTTGGTAAGCTTGGCTTGG 156
 Db 10769 GAAATGAGTATGCTTAGGGGATTTTAAGGATTTTATTGTTTTTGGAAATTTGGTGG 10828
 QY 157 CTTCCTGGTAAATATGAGAAGTAAGTAAGTAAAGTATTTGACTCTAGTCAAGTACATGG 216
 Db 10829 TTTATTTAGTACGAGTATGATGCTTAAAGTATTTTATAGTTTGAAGTATTTT 10888
 QY 217 GATTGCCCTTTTCGGGCTTGGATGGCTTGGTGGTGGTGGAGCAACAATTTATAG 276
 Db 10889 ATTTTCTGTTGGTTAGATTGAGTTTATTTATTTTGT---TTTTTTTTTTTATTA 10945
 QY 277 AAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 336
 Db 10946 AAATTTAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11005
 QY 337 AATATACACAGAGAGTACTTT 358
 Db 11006 AAAAAATTACGTTGTATGATTTT 11027

RESULT 11
 AAS46643
 ID AAS46643 standard; DNA; 6261 BP.
 AC AAS46643;
 XX
 DT 18-DEC-2001 (first entry)
 DE
 XX Tumour suppressor gene derived chemically modified sequence #365.
 XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
 KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 XX Homo sapiens.
 XX
 XX WO200168912-A2.
 FN
 XX
 PD 20-SEP-2001.
 XX
 PF 15-MAR-2001; 2001WO-EP02955.
 XX
 XX 15-MAR-2000; 2000DE-1013847.
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX

PI Olek A, Piepenbrock C, Berlin K;
 XX WPI; 2001-602752/68.
 DR
 XX Fragments of chemically modified genes associated with tumour suppressor
 PT genes and oncogenes, useful in designing primers and probes for
 PT analysing diseases associated with cytosine methylation state e.g.
 PT cancer
 XX
 XX Claim 1; SEQ ID No 365; 27pp; English.
 PS The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and
 CC oncogenes having a sequence taken from 536 (actually 533 since
 CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
 CC (SS) and sequences complementary to (SS). The nucleic acid may be a
 CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
 CC form part of a set of probes for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the
 CC differences serving as basis for diagnosis and/or prognosis events which
 CC are disadvantageous to patients. The present sequence is one of the
 CC 533 genomic sequences derived from tumour suppressor genes and
 CC oncogenes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 6261 BP; 1716 A; 190 C; 1696 G; 2659 T; 0 other;
 Query Match 3.3%; Score 44; DB 22; Length 6261;
 Best Local Similarity 50.5%; Pred. No. 0.35;
 Matches 107; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
 QY 146 GGTTCGGCTTGGCTTCTGGGTAATATGAGAAGTAAGTAAGTAAGTAAGTAAGTAAGT 205
 Db 1216 GGTCTGTTGTAGTTTGTAGTTTTCGGGAGGTTGAGTAGGAGATGCTGTGAATTCGGGA 1275
 QY 206 CAAGTACATTCGATTCCTTTTTCGGGCTTGGATTCGGTTCGGTTCGGAGCCAA 265
 Db 1276 GCGCGAGTTTGTAGTTCGAGTCGAGATCGCGTTATTGTATTTCGGGCGATAGAGCGA 1335
 QY 266 CAATTTATAAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 325
 Db 1336 GATTTTATATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1395
 QY 326 GGTGGGCGAGAAATATACACAGAGAGTACTTT 357
 Db 1396 AGGTAAGTAAATTTTATAACGAAGATATT 1427

RESULT 12
 AAS61046
 ID AAS61046 standard; DNA; 6261 BP.
 XX
 AC AAS61046;
 XX
 DT 29-JAN-2002 (first entry)
 DE
 XX Human gene regulation-associated gene oligonucleotide #1.
 DE Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preeclampsia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;

	Matches	92;	Conservative	0;	Mismatches	82;	Indels	0;	Gaps	0;
Qy	151	GCTTGGCTTCTGCGTAATATGAGAGTAAAGAGTAAAGGTATTGACCTCTAGTCAAGT	210							
Db	164	GTITGGGTTTGGGTTAGAGAGAGTGAATGAAATAGAGAGATTGTATTAACTAGCCTAGG	223							
Qy	211	ACATTGGATTGCCCTTCTCGGGGCTTGGATGGCTTGGTTTCGTGTGAGAGCCCAACAATT	270							
Db	224	AAAATGGTGCACCTGAATAAAGGTTACTGCGAGGGTTTCCCCCATGCAAGCANAAAAA	283							
Qy	271	TATAGAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	324							
Db	284	AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	337							

Search completed: January 21, 2004, 17:31:46
 Job time : 344.379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 17:02:36 ; Search time 2632.89 Seconds
(without alignments)
12185.050 Million cell updates/sec

Title: US-09-936-869-1

Perfect score: 1320

Sequence: 1 aaactgcagcaaatatga.....tggtgctgtgcgcattg 1320

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_estum.*

3: em_estin.*

4: em_estmu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_estc.*

9: gb_estc.*

10: gb_estc.*

11: gb_htc.*

12: gb_estc.*

13: gb_estc.*

14: gb_estc.*

15: em_estfun.*

16: em_estom.*

17: em_gss_hum.*

18: em_gss_inv.*

19: em_gss_pln.*

20: em_gss_vrt.*

21: em_gss_fun.*

22: em_gss_lman.*

23: em_gss_mus.*

24: em_gss_pro.*

25: em_gss_rod.*

26: em_gss_phg.*

27: em_gss_vri.*

28: gb_gss.*

29: gb_gss.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	154	11.7	526	9 AW754659	AW754659 PC04G10 P
2	61.4	4.7	1101	29 CNS0039G	AL063921 Drosophil
3	57.2	4.3	987	23 CNS00418	AL06537 Drosophil
4	56.2	4.3	1201	13 BX461310	BX461310 BX461310

5	55.2	4.2	1101	29 CNS003B2	AL064080 Drosophil
6	55.2	4.2	1101	29 CNS00DG1	AL069971 Drosophil
7	54.8	4.2	1041	9 AL559034	AL559034 AL559034
8	54	4.1	964	29 CNS006N9	AL065781 Drosophil
9	54	4.1	1101	29 CNS0174K	AL107582 Drosophil
10	54	4.1	1101	29 CNS0170K	AL108152 Drosophil
11	53.4	4.0	1101	29 CNS0100X	AL098379 Drosophil
12	52.8	4.0	997	29 CNS005TE	AL060767 Drosophil
13	52.6	4.0	634	29 BX122696	AL122696 Dario rer
14	52.6	4.0	1114	13 CNS0182P	AL108811 Drosophil
15	52.6	4.0	1101	29 BX407949	AL108811 Drosophil
16	52.6	4.0	1201	9 AL565656	AL565656 AL565656
17	52	3.9	783	29 CNS00A1S	AL055833 Drosophil
18	52	3.9	797	13 BX437575	AL055833 Drosophil
19	52	3.9	1101	29 CNS0181N	AL108773 Drosophil
20	51.8	3.9	623	28 AQ157491	AQ157491 nbxb0009H
21	51.8	3.9	830	29 CNS06P72	AL409029 T7 end of
22	51.8	3.9	1099	13 BX456575	AL456575 BX456575
23	51.6	3.9	946	13 BX416277	AL416277 BX416277
24	51.6	3.9	1201	13 BX461128	AL461128 BX461128
25	51.4	3.9	1080	28 AQ738730	AQ738730 HS_5382 B
26	51.4	3.9	1101	29 CNS0039R	AL063932 Drosophil
27	51.2	3.9	997	29 CNS005TE	AL060767 Drosophil
28	51.2	3.9	1204	29 CNS016E2	AL106628 Drosophil
29	51	3.9	673	29 CNS03XD5	AL264866 Tetraodon
30	50.8	3.8	1011	13 BX349844	AL349844 BX349844
31	50.6	3.8	929	13 BX408907	AL408907 BX408907
32	50.6	3.8	938	29 CNS006TJ	AL065906 Drosophil
33	50.4	3.8	626	28 AZ020346	AZ020346 RPL-23-3
34	50.4	3.8	924	29 CNS01GHN	AL143148 Anopheles
35	50.4	3.8	1001	29 CNS006G4	AL062781 Drosophil
36	50.2	3.8	1101	29 CNS017KK	AL108171 Drosophil
37	50	3.8	1049	13 BX353130	AL353130 BX353130
38	50	3.8	1101	29 CNS0106X	AL098595 Drosophil
39	50	3.8	1201	13 BX335607	AL335607 BX335607
40	50	3.8	1201	13 BX358867	AL358867 BX358867
41	49.6	3.8	658	28 AZ851742	AZ851742 ZMO154E07
42	49.6	3.8	735	29 AG095704	AG095704 Pan trogl
43	49.6	3.8	1201	13 BX460090	AL460090 BX460090
44	49.2	3.7	732	28 BZ146517	BZ146517 CH230-453
45	49.2	3.7	760	29 AG056350	AG056350 Pan trogl

ALIGNMENTS

RESULT 1
AW754659

LOCUS

DEFINITION

PC04G10 Pine Triplex pollen cone library Pinus taeda CDNA clone

ACCESSION

AW754659

VERSION

AW754659.1 GI:7676379

KEYWORDS

EST.

SOURCE

Pinus taeda (loblolly pine)

ORGANISM

Pinus taeda

REFERENCE

Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

1 (bases 1 to 526)

Db 952 KAGWAGAAARAGARAAAAARRAGAGANTGAAAGGTGGGAARGGATGTGTAAGRG 893
Qy 233 GCTTGGATGCTTGGTTCGTGTGAGAGCCAAATTTATAGAGAAATATATAAAATAAA 292
Db 892 GGATRAAAATTTTWSAAAAVDTSSTSAASAAAARAAVAAAAAARAAAAA 833
Qy 293 AATAAAAAATTTAGTGTGGAAGTGAAGAACGGTGGGCGCAAAATATACAGAGAG 352
Db 832 WAAAAAATWAAAAASTSTAAAAASSAVAAASAWWWATGSSRTAAAWARSARGA 773
Qy 353 TACTTTAACCAATGGCAACCAAGCAGATTCACAACTTGATTTCTGGACCTCGAATACGA 412
Db 772 AARTWAAAAAAGAAAAAASSVASASAWRAVSSASAAAWAAARVAAAT 713
Qy 413 GATAATGCTGTGAAGAAATAAGAGAGAGTGGAGTGCATTGAAATGAATGAGAGCGC 472
Db 712 GATAAGGTGGGAGAAARTTTWAATTAAGWRRRAATGKTAWRARAATATGAARSAG 653
Qy 473 ACAAAATGGAGGAGCAATTAATGAATATATAT 504
Db 652 ARTWAATWRAAATGWAATAATTTARTTTGAT 621

RESULT 7
AL559034/c
LOCUS
DEFINITION
AL559034 Homo sapiens T CELLS (JURKAT CELL LINE) EST 31-MAY-2003
Homo sapiens cDNA clone CS0DJ010YL15 5-PRIME, mRNA sequence.
ACCESSION
AL559034
VERSION
AL559034.2 GI:31283167
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 15, 2001 this sequence version replaced gi:12904134.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by life technologies, a
division of invitrogen. This sequence belongs to sequence cluster
8934.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DJ010CF08QPI&cluster=8934.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/InvitroGen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DJ010CF08QPI.
Location/Qualifiers
1. .1041
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ010YL15"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT
153 a 190 c 201 g 335 t 162 others
ORIGIN

Query Match 4.2%; Score 54.8; DB 9; Length 1041;
Best Local Similarity 33.7%; Pred. No. 26;
Matches 86; Conservative 29; Mismatches 140; Indels 0; Gaps 0;

Qy 254 GTGGAAGCCCAACATTTATAGAAATATATAAAATAAAAATTTAAAGTGT 313
Db 460 GAGTWWAAAAARGCCTGWRAAAAAARAAATATAAGAAATGAAAAARWAAAAAGAA 401
Qy 314 GGAAGTGAACCGTGGGCGAGAAATATACACAGAGTACTTTTAAACAATGCCCAACA 373
Db 400 AAAAAAAMAAWAAWRAARRAARRAAAAAARAAAAAARAAAAAARAAAAAARAAAAA 341
Qy 374 AGGCAGATTCACAACTTGATTTCTGGACCTCGAATACAGATAATGTGGTAAAGAAATAA 433
Db 340 NAAAAAANNA 281
Qy 434 AGGAGAGTGGAGTGCATTGAAAAATGAATGAGAGCGCACAAAATGGAGACGAAATAA 493
Db 280 ANNAANNN 221
Qy 494 TCAATAATAATGCAA 508
Db 220 RCARATCAACGARM 206

RESULT 8
CNS006N9
LOCUS
DEFINITION
CNS006N9 Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14719 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL065781
VERSION
AL065781.1 GI:4944661
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 964)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source
1. .964
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14J19"
/clone_lib="RPCI-98"
/note="end : T7"
BASE COUNT
409 a 111 c 71 g 130 t 243 others
ORIGIN

Query Match 4.1%; Score 54; DB 29; Length 964;
Best Local Similarity 34.0%; Pred. No. 36;
Matches 86; Conservative 64; Mismatches 103; Indels 0; Gaps 0;

Qy 264 AACCAATTTAAGAAATATATAAAATAAAAATTTAAAGTGTGGAGTGA 323
Db 264 AACCAATTTAAGAAATATATAAAATAAAAATTTAAAGTGTGGAGTGA 323

```

Db      664 AAAAAATTGTTAGTAAATATAAAAAATAAAAAADADAADAAAKKTKWDKADAAAAA 723
QY      324 ACCGTGGGCGAGAAATATACACAGAAGAGTACTTTAACCAATGCGCAACCAAGCAGATTC 383
Db      724 AAGWAKKDAAAAAAATAAAAAAATAAAAAAATADADAADATAKAKWDWADAKDDDKAGA 783
QY      384 ACAACTTGATTCTGACCTCGAATACGAGATTAATGGTGTGAAGAAATAAAGGAAGGTG 443
Db      784 AKADKARKAARAAAANKCKDKARKAKAWAKAKAGTGWAAKARAGKATAGAGAGAK 843
QY      444 GAGTGCATTTGAAAAATGAATGGAGAGCGGCACAAAATCGAGGACGAGTAATAATGAATATAA 503
Db      844 AAKAKDKAGKADAKADAKADAKADAKADAKADAKADAKADAKADAKADAKADAKADAK 903
QY      504 TGCAAGAGTGCAT 516
Db      904 AKAAAAAARWAK 916

RESULT 9
CNS0174K      1101 bp      DNA      linear      GSS 26-JUL-1999
LOCUS      Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION      BACN17P20 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION      AL107582
VERSION      AL107582.1 GI:5627886
KEYWORDS
SOURCE      GSS.
ORGANISM      Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 1101)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY Cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT      - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -- This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.
FEATURES
    source      Location/Qualifiers
    1..1101
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /db_xref="taxon:7227"
    /clone="BACN17P20"
    /clone_lib="DrosBAC"
    /plasmid="pBelobAC11"
    /note="end : SP6"

BASE COUNT      365 a      90 c      235 g      206 t      205 others
ORIGIN
    Query Match      4.1%; Score 54; DB 29; Length 1101;
    Best Local Similarity 42.5%; Pred. No. 35;
    Matches 147; Conservative 22; Mismatches 177; Indels 0; Gaps 0;

QY      253 TGTGAGAGCCAACTTATAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAA 312
Db      216 TATAATAGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 275
QY      313 TGGAGTGAACCGTGGGCGAGAAATATACACAGAAGAGTACTTTAACCAATGCGCAACC 372
Db      276 AATTAAAAAARWAGWAGAAAAATATATAAATAAAGTARWTATTAATTAATGAAT 335
QY      373 AAGCAGATTCACACTTGATTTCTGGACCTCGAATACGAGATTAATGGTGTGAAGATA 432

```

```

Db      336 AAWTTATATWAAAAARGAAAGATGAWAGRAAAAAATRAAAAAAATAAATAAATAA 395
QY      433 AAGGAAGAGTGGAGTGCATTGCAATTTGAAAAATGAATGGAGAGCGCACAAAATGGAGCAGATAA 492
Db      396 AWTWAAAAAAMWATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 455
QY      493 ATGAAATATATGCAAGAGTGCATTTCCCTATTATTATTTCCAGAAATGTATATGTGGGGTCG 552
Db      456 AAGTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 515
QY      553 GCATTCACATGGGCGTGCATTCAGGGGGTGTGCATACGGGTCCCTTT 598
Db      516 GCATTCAMGAASGTGCGKAAAWAGCGGTACCGGTATCSTATTCTT 561

RESULT 10
CNS017KE/c      1101 bp      DNA      linear      GSS 26-JUL-1999
LOCUS      Drosophila melanogaster genome survey sequence T7 end of BAC
DEFINITION      BACN37F11 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION      AL108152
VERSION      AL108152.1 GI:5628456
KEYWORDS
SOURCE      GSS.
ORGANISM      Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 1101)
AUTHORS      Genoscope.
TITLE      Direct Submission
JOURNAL      Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT      - Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk -- This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelobAC11.
FEATURES
    source      Location/Qualifiers
    1..1101
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /db_xref="taxon:7227"
    /clone="BACN37F11"
    /clone_lib="DrosBAC"
    /plasmid="pBelobAC11"
    /note="end : T7"

BASE COUNT      302 a      83 c      32 g      379 t      305 others
ORIGIN
    Query Match      4.1%; Score 54; DB 29; Length 1101;
    Best Local Similarity 35.1%; Pred. No. 35;
    Matches 99; Conservative 57; Mismatches 126; Indels 0; Gaps 0;

QY      264 AACAAATTTAAGAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 323
Db      711 MAGDAAKKGAAAAAHHADAAAAAAGGWAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHA 652
QY      324 ACGGTGGGCGAGAAATATACACAGAAGAGTACTTTAACCAATGCGCAACCAAGCAGATTC 383
Db      651 AAKGAAGTTCAGACACAKAKAAAGAGDCKKGGKDKAKAAAKAKAAAKAKAKAKAKAK 592
QY      384 ACAACTTGATTCTGACCTCGAATACGAGATTAATGGTGTGAAGAAATAAAGGAAGGTG 443
Db      591 KKKAKKAAAKKDKWCRAKKWAAAKDKKAGAGAADKKGAAGAAAGAAKAAATATATWGA 532
QY      444 GAGTGCATTTGAAAAATGAATGGAGAGCGCACAAAATGAGGACGAGTAATAATGAATATAA 503

```


FEATURES
source

BASE COUNT
ORIGIN

Search completed: January 22, 2004, 00:54:15
Job time : 2640.89 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 16:52:57 ; Search time 89.2891 Seconds
(without alignments)
6525.161 Million cell updates/sec

Title: US-09-936-869-1
Perfect score: 1320
Sequence: 1 aaactcgagcaaatatga.....tattgctgtggccattg 1320

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCUS COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	73.4	5.6	7218	1	US-08-232-463-14
C 2	61	4.6	7218	1	US-08-232-463-14
C 3	43.6	3.3	176373	3	US-08-128-155-17
C 4	42	3.2	152331	3	US-08-128-155-17
C 5	41	3.1	2255	2	US-08-741-134-1
C 6	40	3.0	87350	3	US-08-781-891-79
C 7	40	3.0	87350	4	US-08-618-166-79
C 8	40	3.0	87543	4	US-08-791-211-3
C 9	39.4	3.0	435	4	US-08-397-787-171
C 10	39.4	3.0	1024	4	US-08-328-475C-64
C 11	39.2	3.0	1753	3	US-08-058-260-21
C 12	39.2	3.0	1756	3	US-08-058-260-31
C 13	39.2	3.0	1776	3	US-08-058-260-23
C 14	39.2	3.0	1952	3	US-08-058-260-5
C 15	39.2	3.0	4090	2	US-08-781-802-5
C 16	39.2	3.0	4090	3	US-08-694-078-5
C 17	38.8	2.9	3564	2	US-07-594-921C-9
C 18	38.8	2.9	3564	4	US-09-197-948-9
C 19	38.6	2.9	202001	4	US-08-734-674-3
C 20	38.2	2.9	289	3	US-09-007-005-17
C 21	38.2	2.9	289	3	US-08-244-796-17
C 22	37.6	2.8	1896	3	US-09-058-260-3
C 23	37.6	2.8	6263	2	US-08-781-802-3
C 24	37.6	2.8	6263	3	US-08-694-078-3
C 25	37.4	2.8	475	4	US-09-702-705-36
C 26	37.4	2.8	475	4	US-09-736-457-36
C 27	37.4	2.8	4235	4	US-09-702-705-317

C 28	37.4	2.8	4235	4	US-09-736-457-317
C 29	37.4	2.8	72604	4	US-08-288-992-7
C 30	37.4	2.8	72604	4	US-08-657-474-7
C 31	36.8	2.8	289	3	US-09-007-005-17
C 32	36.8	2.8	289	3	US-09-244-796-17
C 33	36.8	2.8	246240	2	US-08-724-394A-20
C 34	36.8	2.8	246240	2	US-08-724-394A-21
C 35	36.8	2.8	246240	2	US-08-724-394A-22
C 36	36.6	2.8	1664976	4	US-08-916-421B-1
C 37	36.4	2.8	658	3	US-08-998-416-595
C 38	36.4	2.8	1164	4	US-09-538-871-3
C 39	36.4	2.8	1232	4	US-09-538-871-1
C 40	36.2	2.7	2013	4	US-09-134-001C-103
C 41	36.2	2.7	176373	3	US-09-138-155-17
C 42	36	2.7	2518	3	US-09-433-699-3
C 43	36	2.7	1230025	4	US-09-198-452A-1
C 44	35.8	2.7	1141	2	US-08-323-449B-1
C 45	35.8	2.7	1141	2	US-08-485-981-1

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0239
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)836-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
; US-08-232-463-14

Query Match 5.6%; Score 73.4; DB 1; Length 7218;

```
Best Local Similarity 6.2%; Pred. No. 9.9e-11;
Matches 26; Conservative 236; Mismatches 157; Indels 0; Gaps 0;
QY 93 CACTGTACTGTAGGAAAGTAGTGGACTCTGCTCAGGTACATTTAGTTGGTAAGTTGGC 152
D 1487 CACTGTAATACCTATCTATGCAAGTAGTTAAAGAGATAGAGAATTTGGTACRRRRR 1428
QY 153 TTGGCTTCCTGGTAAATATGAGAGTAAAGCAAGTAAAGGTATTTCACTCTAGTCAAGTAC 212
D 1427 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1368
QY 213 ATTGGATGCTTTGTGCGGGCTTGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 272
D 1367 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1308
QY 273 TAAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 332
D 1307 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1248
QY 333 CAGAAATATACAGAGAGAGTACTTTAAATGCGCAACCAAGGAGATTCACAACTTGA 392
D 1247 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1188
QY 393 TTCTCGACTCGAATACAGATATAATGTTGGTAAAGATAAAGAGAGTGGAGTGCATT 452
D 1187 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1128
QY 453 TGAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 511
D 1127 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1069

RESULT 2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESS: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F18
; US-08-232-463-14
```

```
Query Match
Best Local Similarity 4.6%; Score 61; DB 1; Length 7218;
Matches 7; Conservative 200; Mismatches 110; Indels 0; Gaps 0;
QY 603 GCAGTGTGGAGTTGCAACATGTACCAACAAATTCATTTCATCCAAACACCAATTTATC 662
D 1055 GGAGCTTCGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1114
QY 663 CTCCTCATCTATCTACCTACATCTACCTATACCTAGTAATATGTCTGCTGTGAACCTCT 722
D 1115 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1174
QY 723 CACTGCTGCACAGCTCTTAGTCAATCCATCTGCTTCATTAATAGGCATTTATTTGTTCTT 782
D 1175 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1234
QY 783 TCCCTCCGACTGAAGGCTATCGACCGACGCGCTCATCTCTCTCTCTCGCAATT 842
D 1235 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1294
QY 843 TTTTGTGCTGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATG 902
D 1295 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1354
QY 903 TATCTCATCTCTCCCT 919
D 1355 YYYYYYYYYYYYYYYYYY 1371
```

```
RESULT 3
US-09-128-155-17/c
; Sequence 17, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Fan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; EARLIER FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17
```

```
Query Match
Best Local Similarity 3.3%; Score 43.6; DB 3; Length 176373;
Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 274 AAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 333
D 122513 AAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 122454
QY 334 AGAAATATACACAGAGAGTACTTTTACAAATGCGCAACCAAGCAGATTTCACAACTT 393
```


ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER: IBM PC compatible
MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800

ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419

TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 87350 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-781-891-79

Query Match 3.0%; Score 40; DB 3; Length 87350;
Best Local Similarity 50.8%; Pred. No. 1;
Matches 120; Conservative 0; Mismatches 115; Indels 1; Gaps 1;

QY 269 TTTATAGAAATATATAAAATAAAATAAAATTTAAAGTGTGGAAGTGAACCGGT 328

Db 49802 TTGGTAAATATATAATATGATATATATAATATAAAAGAAATTTTAAATCTAAACCGT 49743

QY 329 GGGGCGAGAAATATACAGAGAGTACTTTAA-CAATGCCAACCGAGGCGAGATTCCAA 387

Db 49742 TTTAAATAATTTTAAATAAGTGAATCTAAGTGTGGAATACAGAAATAAATACAA 49683

QY 388 CTTCATTTCTGACCTCGAATACGAGATAATGTTGTAAGAAATAAAGGAGAGTGGAGT 447

Db 49682 TTATGTTAAAGAGAAATAGAGGTGAAAGGAGAAATAAATTTAAAGAGAAATGAAGA 49623

QY 448 GCATTTCAAATGAATGAGAGCGCACAAAATGAGGACGCAATAAATGAATATAA 503

Db 49622 AATAAAGAGATGATGAAAGTGAATAATTTGAAGCTGGTAAATGAAGATCCAA 49567

RESULT 7

US-09-618-166-79/c
Sequence 79, Application US/09618166
Patent No. 6583112

GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
Yu, Chang-En

Oshima, Junko
Mulligan, John T.
Schellenberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
WERNER'S SYNDROME

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/618,166

FILING DATE: 17-JUL-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: McMasters, David D.

REGISTRATION NUMBER: 33,963

REFERENCE/DOCKET NUMBER: 240052.419C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 79:

SEQUENCE CHARACTERISTICS:

LENGTH: 87350 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 79:

US-09-618-166-79

Query Match 3.0%; Score 40; DB 4; Length 87350;
Best Local Similarity 50.8%; Pred. No. 1;
Matches 120; Conservative 0; Mismatches 115; Indels 1; Gaps 1;

QY 269 TTTATAGAAATATATAAAATAAAATAAAATTTAAAGTGTGGAAGTGAACCGGT 328

Db 49802 TTGGTAAATATATAATATGATATATAATATAAAAGAAATTTTAAATCTAAACCGT 49743

QY 329 GGGGCGAGAAATATACAGAGAGTACTTTAA-CAATGCCAACCGAGGCGAGATTCCAA 387

Db 49742 TTTAAATAATTTTAAATAAGTGAATCTAAGTGTGGAATACAGAAATAAATACAA 49683

QY 388 CTTCATTTCTGACCTCGAATACGAGATAATGTTGTTAAGAAATAAAGGAGAGTGGAGT 447

Db 49682 TTATGTTAAAGAGAAATAGAGGTGAAAGGAGAAATAAATTTAAAGAGAAATGAAGA 49623

QY 448 GCATTTGAATGAATGAGAGCGCACAAAATGAGGACGCAATAAATGAATATAA 503

Db 49622 AATAAAGAGATGATGAAAGTGAATAATTTGAAGCTGGTAAATGAAGATCCAA 49567

RESULT 8

US-09-791-211-3/c
Sequence 3, Application US/09791211
Patent No. 6448080

GENERAL INFORMATION:
APPLICANT: Donna T. Watt

APPLICANT: Andrew T. Watt

TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION

FILE REFERENCE: RTS-0205

CURRENT APPLICATION NUMBER: US/09/791,211

CURRENT FILING DATE: 2001-02-23

NUMBER OF SEQ ID NOS: 90

SEQ ID NO 3

LENGTH: 87543

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: unsure

LOCATION: 7421

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 7427

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 11609

OTHER INFORMATION: unknown

NAME/KEY: unsure

LOCATION: 12605

OTHER INFORMATION: unknown

NAME/KEY: unsure
LOCATION: 12742
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29370
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29422
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29979
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29980
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29981
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30136
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30140
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31205
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31206
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31592
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 33095
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 33160
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34066
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34072
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 36916
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 39020
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 42164
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 42459
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46808
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46823
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46826
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 47291
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 52786
OTHER INFORMATION: unknown
NAME/KEY: unsure

LOCATION: 52787
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 53384
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 54684
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59215
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59235
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59242
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 63290
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 66614
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68660
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68697
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68718
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68733
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68739
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 69785
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79134
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79198
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 86336
OTHER INFORMATION: unknown
OTHER INFORMATION:
US-09-791-211-3

Query Match 3.0%; Score 40; DB 4; Length 87543;
Best Local Similarity 50.8%; Pred. No. 1;
Matches 120; Conservative 0; Mismatches 115; Indels 1; Gaps 1;
QY 269 TTTATAGAAATATATAAAATAAAAAATTTTAACTGTGGAGTGAACGGT 328
Db 49995 TTGGTAAAAATAAATATGTATATATAATAAAAAAGAAATTTTAAATCTAAAAAACGT 49936
QY 329 GGGGAGAAATATACACAGAGAGTACTTTAA-CAATGCGCAACCAAGCGAGATTCAAA 387
Db 49935 TTTAAAAATATTTTAAATTAAGTGAATCTAGTGTGGAATACAGAAATAATACAA 49876
QY 388 CTTGATTTCTGACCTCGAATACGAGATAATGGTGGTAAAGAAATAAGAGAGAGTGGAGT 447
Db 49875 TTATGTTAAAAAGAGAAATAGAAAGGTGAAAGAGGAAAAAATTTAAAGAGAAATGAAGA 49816
QY 448 GCATTTGAAAAATGATGAGAGCGCACAAATGGAGCGAATAAATGAAATATAA 503
Db 49815 AATAAAAGAAATGCATGAAAGTGAATAATTTGAGCTGGTAAATGAAGATCCAA 49750

```
RESULT 9
US-09-397-787-171/C
; Sequence 171, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Benson, Michael J.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 171
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-171

Query Match      3.0%; Score 39.4; DB 4; Length 435;
Best Local Similarity 71.2%; Pred. No. 0.17; 21; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 0;

QY 825 TTCTCTCTCGGCAATTTTCTGCTGGATCATCATATACATCATCGGCATCCCA 884
DB 372 TTCTTCACTGGCGCTTTTCTCAGCTTCCTCATCATCAAAATCATCATCATCT 313

QY 885 CCATCATCATCAT 897
DB 312 TCATCATCATCCT 300

RESULT 10
US-09-328-475C-64/C
; Sequence 64, Application US/09328475C
; Patent No. 6476207
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steinmann, Kathleen E.
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
; FILE REFERENCE: 1532.002/200130.463
; CURRENT APPLICATION NUMBER: US/09/328,475C
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo Sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1024)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-64

Query Match      3.0%; Score 39.4; DB 4; Length 1024;
Best Local Similarity 71.2%; Pred. No. 0.17; 21; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 0;

QY 825 TTCTCTCTCGGCAATTTTCTGCTGGATCATCATATACATCATCGGCATCCCA 884
DB 440 TTCTTCACTGGCGCTTTTCTCAGCTTCCTCATCATCAAAATCATCATCATCT 381
```

```
QY 885 CCATCATCATCAT 897
DB 380 TCATCATCATCCT 368

RESULT 11
US-09-058-260-21
; Sequence 21, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fonstein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demirjian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; CURRENT FILING DATE: 1993-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1753
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cloned esterase
; OTHER INFORMATION: gene from bacteria E013
; NAME/KEY: CDS
; LOCATION: (128)..(1630)
US-09-058-260-21

Query Match      3.0%; Score 39.2; DB 3; Length 1753;
Best Local Similarity 56.1%; Pred. No. 0.26; 58; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 58;

QY 205 TCAAGTACATTGGATTGCTTTGTGGGGCTTGGATGGCTTGGGTCTGTGAGAGCCA 264
DB 24 TCCGGTGGCATGGATTGCTTCAGGGGAACCTTTAAACACTTGAGTTTGACACCACTCCT 83

QY 265 ACAATTTATAGAAATATATAATAATAATAATAATAATAATAATAATAATAATA 324
DB 84 TAATCATTTAAGATTAAATGAAATTAATAATAATAATAATAATAATAATAATAATA 143

QY 325 CGGTGGGGCAGA 336
DB 144 CGTTGGTGGAAA 155

RESULT 12
US-09-058-260-31
; Sequence 31, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fonstein, Michael
; APPLICANT: Vonstein, Veronika
```

```
; APPLICANT: Demirjian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; EARLIER FILING DATE: 1999-04-10
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 1756
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; OTHER INFORMATION: gene from bacteria E027
; NAME/KEY: CDS
; LOCATION: (130)..(1632)
; US-09-058-260-31

Query Match          3.0%; Score 39.2; DB 3; Length 1756;
Best Local Similarity 56.1%; Pred. No. 0.26;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 205 TCAAGTACATGGATTGCTTTGTCGGGGCTTGGATGCTTGGTTCGTGTGAGAGCCA 264
Db 26 TCCGGTGCATGGATTGCTTCAGGGGAACCTTTAAACACTTGAGTTGACACCACTCCT 85
QY 265 ACAATTTATAAGAAATATATAAAATAAAAAATTAAGTTGTTGGAAGTGAATA 324
Db 86 TAATCATTTAAGATTAAATGAAATTAATAATCAAAAAGAGTGATTCAAATGAATA 145
QY 325 CGGTGGGCAGCA 336
Db 146 CGTTGGTGGAAA 157

RESULT 13
US-09-058-260-23
; Sequence 23, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fomstein, Michael
; APPLICANT: Vonstein, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; EARLIER FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 1756
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; OTHER INFORMATION: gene from bacteria E011
; NAME/KEY: CDS
; LOCATION: (197)..(1699)
```

```
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1776
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; OTHER INFORMATION: gene from bacteria E015
; NAME/KEY: CDS
; LOCATION: (128)..(1630)
; US-09-058-260-23

Query Match          3.0%; Score 39.2; DB 3; Length 1776;
Best Local Similarity 56.1%; Pred. No. 0.26;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 205 TCAAGTACATGGATTGCTTTGTCGGGGCTTGGATGCTTGGTTCGTGTGAGAGCCA 264
Db 24 TCCGGTGCATGGATTGCTTCAGGGGAACCTTTAAACACTTGAGTTGACACCACTCCT 83
QY 265 ACAATTTATAAGAAATATATAAAATAAAAAATTAAGTTGTTGGAAGTGAATA 324
Db 84 TAATCATTTAAGATTAAATGAAATTAATAATCAAAAAGAGTGATTCAAATGAATA 143
QY 325 CGGTGGGCAGCA 336
Db 144 CGTTGGTGGAAA 155

RESULT 14
US-09-058-260-5
; Sequence 5, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fomstein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demirjian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; EARLIER FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1952
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; OTHER INFORMATION: gene from bacteria E011
; NAME/KEY: CDS
; LOCATION: (197)..(1699)
```


US-09-058-260-5

Query Match 3.0%; Score 39.2; DB 3; Length 1952;
Best Local Similarity 56.1%; Pred. No. 0.27;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 205 TCAAGTACATTTGGATTCCTTTTCGGGGCTTGGGTTTCGGTTCGTGTGAGAAGCCA 264
DB 93 TCCGGTGCATGGATTGCTTCAGGGGAACCTTTTAAACACTTGAGTTTGACAAACCCTCCT 152
QY 265 ACAATTTAAGAATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 324
DB 153 TAATCATTTAAGATTAAATGAATAAATAAATAAATAAATAAATAAATAAATAAATAA 212
QY 325 CGGTGGGCAG 336
DB 213 CGTGTGGGAAA 224

RESULT 15

US-08-781-802-5
; Sequence 5, Application US/08781802
; Patent No. 5969121
; GENERAL INFORMATION:
; APPLICANT: ALLEN, Larry
; APPLICANT: AIRENS, John
; APPLICANT: FORSTEIN, Michael
; APPLICANT: VONSTEIN, Veronika
; APPLICANT: DEMIRJIAN, David
; APPLICANT: CASADABAN, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,802
; FILING DATE: 10-JAN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/694,078
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/019,580
; FILING DATE: 12-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/009,704
; FILING DATE: 11-JAN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/001,995
; FILING DATE: 01-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Chao, Mark
; REGISTRATION NUMBER: 37,293
; REFERENCE/DOCKET NUMBER: 95,963-E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4090 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 197..1699
; OTHER INFORMATION: /note= "E011 sequence of longest
; OTHER INFORMATION: open reading frame; other possible start codons ATG/met4;
; OTHER INFORMATION: TTG/1eu7; GTG/val15; GTG/val36; ATG/met62"
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 197..1699
; US-08-781-802-5

Query Match 3.0%; Score 39.2; DB 2; Length 4090;
Best Local Similarity 56.1%; Pred. No. 0.39;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 205 TCAAGTACATTTGGATTCCTTTTCGGGGCTTGGGTTTCGTGTGAGAAGCCA 264
DB 93 TCCGGTGCATGGATTGCTTCAGGGGAACCTTTTAAACACTTGAGTTTGACAAACCCTCCT 152
QY 265 ACAATTTAAGAATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 324
DB 153 TAATCATTTAAGATTAAATGAATAAATAAATAAATAAATAAATAAATAAATAAATAA 212
QY 325 CGGTGGGCAG 336
DB 213 CGTGTGGGAAA 224

Search completed: January 21, 2004, 22:19:40
Job time : 94.2891 secs

Sequence 2003, Ap
Sequence 2359, Ap
Sequence 1979, Ap
Sequence 1, Appl1
Sequence 18750, A
Sequence 21744, A
Sequence 5016, Ap
Sequence 2003, Ap
Sequence 8, Appl1
Sequence 573, App
Sequence 11962, A
Sequence 2034, Ap
Sequence 8380, Ap
Sequence 1727, Ap
Sequence 3, Appl1
Sequence 4630, Ap
Sequence 1995, Ap
Sequence 185, App
Sequence 2331, Ap
Sequence 1293, Ap
Sequence 3400, Ap
Sequence 82, Appl1
Sequence 9693, Ap
Sequence 16, Appl
Sequence 532, App
Sequence 532, App
Sequence 414, App
Sequence 25620, A
Sequence 8959, Ap
Sequence 340, App

Sequence 2003, Ap
Sequence 2359, Ap
Sequence 1979, Ap
Sequence 1, Appl1
Sequence 18750, A
Sequence 21744, A
Sequence 5016, Ap
Sequence 2003, Ap
Sequence 8, Appl1
Sequence 573, App
Sequence 11962, A
Sequence 2034, Ap
Sequence 8380, Ap
Sequence 1727, Ap
Sequence 3, Appl1
Sequence 4630, Ap
Sequence 1995, Ap
Sequence 185, App
Sequence 2331, Ap
Sequence 1293, Ap
Sequence 3400, Ap
Sequence 82, Appl1
Sequence 9693, Ap
Sequence 16, Appl
Sequence 532, App
Sequence 532, App
Sequence 414, App
Sequence 25620, A
Sequence 8959, Ap
Sequence 340, App

Sequence 2003, Ap
Sequence 2359, Ap
Sequence 1979, Ap
Sequence 1, Appl1
Sequence 18750, A
Sequence 21744, A
Sequence 5016, Ap
Sequence 2003, Ap
Sequence 8, Appl1
Sequence 573, App
Sequence 11962, A
Sequence 2034, Ap
Sequence 8380, Ap
Sequence 1727, Ap
Sequence 3, Appl1
Sequence 4630, Ap
Sequence 1995, Ap
Sequence 185, App
Sequence 2331, Ap
Sequence 1293, Ap
Sequence 3400, Ap
Sequence 82, Appl1
Sequence 9693, Ap
Sequence 16, Appl
Sequence 532, App
Sequence 532, App
Sequence 414, App
Sequence 25620, A
Sequence 8959, Ap
Sequence 340, App

ALIGNMENTS

Sequence 2003, Ap
Sequence 2359, Ap
Sequence 1979, Ap
Sequence 1, Appl1
Sequence 18750, A
Sequence 21744, A
Sequence 5016, Ap
Sequence 2003, Ap
Sequence 8, Appl1
Sequence 573, App
Sequence 11962, A
Sequence 2034, Ap
Sequence 8380, Ap
Sequence 1727, Ap
Sequence 3, Appl1
Sequence 4630, Ap
Sequence 1995, Ap
Sequence 185, App
Sequence 2331, Ap
Sequence 1293, Ap
Sequence 3400, Ap
Sequence 82, Appl1
Sequence 9693, Ap
Sequence 16, Appl
Sequence 532, App
Sequence 532, App
Sequence 414, App
Sequence 25620, A
Sequence 8959, Ap
Sequence 340, App

RESULT 1

Sequence 2003, Ap
Sequence 2359, Ap
Sequence 1979, Ap
Sequence 1, Appl1
Sequence 18750, A
Sequence 21744, A
Sequence 5016, Ap
Sequence 2003, Ap
Sequence 8, Appl1
Sequence 573, App
Sequence 11962, A
Sequence 2034, Ap
Sequence 8380, Ap
Sequence 1727, Ap
Sequence 3, Appl1
Sequence 4630, Ap
Sequence 1995, Ap
Sequence 185, App
Sequence 2331, Ap
Sequence 1293, Ap
Sequence 3400, Ap
Sequence 82, Appl1
Sequence 9693, Ap
Sequence 16, Appl
Sequence 532, App
Sequence 532, App
Sequence 414, App
Sequence 25620, A
Sequence 8959, Ap
Sequence 340, App

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.6	3.8	22684	13	US-10-017-161-725
2	49.6	3.8	29221	12	US-10-232-798-2003
3	45	3.4	519	10	US-09-878-574-4292
4	43.8	3.3	625	13	US-10-027-632-44092
5	43.8	3.3	625	13	US-10-027-632-44092
6	43.8	3.3	625	13	US-10-027-632-44094
7	43.8	3.3	625	13	US-10-027-632-44095
8	43.8	3.3	625	14	US-10-027-632-44092
9	43.8	3.3	625	14	US-10-027-632-44093
10	43.8	3.3	625	14	US-10-027-632-44094
11	43.8	3.3	625	14	US-10-027-632-44095
12	43.6	3.3	26320	13	US-10-034-650-13
13	43.6	3.3	176373	14	US-10-095-407-17
14	43.4	3.3	409	11	US-09-918-995-6799
15	43.4	3.3	13326	13	US-10-311-455-1685

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 17:11:45 ; Search time 444.739 Seconds
(without alignments)
10461.606 Million cell updates/sec

Title: US-09-936-869-1

Perfect score: 1320

Sequence: 1 aaactgcagacaaatatga.....tggtgtgtgtgcgcattg 1320

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*


```
QY 385 CAACCTTGATTTCTGGACCTCGAATACAGATATATGTTGGTAAAGAAATAAAGGAAGAGTGG 444
Db 391 AAAAGTTAAAAATATAATATAAAAAATAAAATTTGAATAAAAAATAATGATAAAATCA 332
QY 445 AGTCATTTGAAATGATGAGGAGCGCCACAAATGAGGAGCAATGAAATGAAATATAAT 504
Db 331 ATTAAAGAAAATATATATAATAATAAAAAATAAAATTAATAATATATATATTAAAGAA 272
QY 505 GCAAGAGTGCATTTCCCTATTATTTCCAGAAAATGTTAT 541
Db 271 TAAAAATTAATGATAAATAATTTGAAAGAAAATAT 235
```

RESULT 4

```
US-10-027-632-44092
; Sequence 44092, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: US 60/127,200
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44092
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(625)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-44092
```

```
Query Match 3.3%; Score 43.8; DB 13; Length 625;
Best Local Similarity 57.8%; Pred. No. 0.31;
Matches 78; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 256 GAGAAGCCCAACAATTATAGAAATATATAAAATAAAATAAAATTTAACTGTTGG 315
Db 349 GAGAGAGACCCCTGCTCTTAATAAAATAAAATAAAATATATAATTAATAATAAAATAAA 408
QY 316 AAGTGAACCAACGGTGGGCGAGAAATATACACAGAGAGTACTTTAACTGCGCAACCAAG 375
Db 409 AAATAAAAGGATAGACAAATAGTCAATTGAAGAGTCCATATAACCAACCAACCAAT 468
QY 376 GCAGATTCACAACTT 390
Db 469 GCAGTCAAAATAGCTT 483
```

RESULT 5

```
US-10-027-632-44093
; Sequence 44093, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44093
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(625)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-44093
```

```
Query Match 3.3%; Score 43.8; DB 13; Length 625;
Best Local Similarity 57.8%; Pred. No. 0.31;
Matches 78; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 256 GAGAAGCCCAACAATTATAGAAATATATAAAATAAAATAAAATTTAACTGTTGG 315
Db 349 GAGAGAGACCCCTGCTCTTAATAAAATAAAATAAAATATATAATTAATAATAAAATAAA 408
QY 316 AAGTGAACCAACGGTGGGCGAGAAATATACACAGAGAGTACTTTAACTGCGCAACCAAG 375
Db 409 AAATAAAAGGATAGACAAATAGTCAATTGAAGAGTCCATATAACCAACCAACCAAT 468
QY 376 GCAGATTCACAACTT 390
Db 469 GCAGTCAAAATAGCTT 483
```

RESULT 6

```
US-10-027-632-44094
; Sequence 44094, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027.632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```



```

; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; PRIOR FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1685
; LENGTH: 13326
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1685

Query Match      3.3%; Score 43.4; DB 13; Length 13326;
Best Local Similarity 53.2%; Pred. No. 2.4; Mismatches 0; Indels 0; Gaps 0;
Matches 92; Conservative 0;

QY 282 TATAAATAAAAAATAAAAAATTTAAGTGTTCGAAAGTGAAAAACGGTGGGCGAGAAATAT 341
Db 9237 TTTTAAATAAGGAGAAATAAATTATATATGTTTCTATGTTAGAGGATAGGAGAGAAATTT 9296

QY 342 ACACAGAGAGAGTACTTTTAAACAATGCCCAACCAAGGAGAGTTCACACTTGATTTCTGGAC 401
Db 9297 TTAGAGAAATAGAAATTTAGGGATGAAGTAGGAGTATAGTTATTAATAAATTAAGTGT 9356

QY 402 CTCGAATACGAGATAATGCGTAAAGAAATAAAGGAAGAGTGGAGTGCATTG 454
Db 9357 ATAGAAGAAAGAAATGTTAGGATGAGATTGGGAAAGTATATTGTGTTTIG 9409

```

Search completed: January 22, 2004, 01:20:32
Job time : 450.739 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 15:50:10 ; Search time 361.266 Seconds
(without alignments)
10468.497 Million cell updates/sec

Title: US-09-936-869-2
Perfect score: 1401
Sequence: 1 aaactgcagcaaatatga.....ggaaattttgtgtacaaatc 1401

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues
Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03: *
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT: *
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT: *
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT: *
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT: *
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT: *
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT: *
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT: *
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT: *
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT: *
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT: *
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT: *
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT: *
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT: *
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT: *
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT: *
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT: *
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT: *
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT: *
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT: *
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT: *
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT: *
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT: *
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT: *
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT: *
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1401	100.0	1401	21	AAA99363
2	1320	94.2	1320	21	AAA99362
3	46.6	3.3	540	24	ABQ18618
4	46.6	3.3	540	24	ABQ18619
5	46.4	3.3	4492	23	ABL15002
6	45.6	3.3	563	23	ABV50724
7	45	3.2	519	25	ABX22233
8	44.4	3.2	15923	24	ABL70179
					Chemically treated
					Plant PRAG1 promot
					Plant PRAG1 promot
					Oligonucleotide fo
					Oligonucleotide fo
					Drosophila melanog
					Human prostate exp
					Human GMP-mannose

9	44.4	3.2	15923	24	AA561132	Human gene regulat
10	44.4	3.2	15923	24	ABK31220	Signal transductio
11	44	3.1	6261	22	AA546643	Tumour suppressor
12	44	3.1	6261	24	AA561046	Human gene regulat
13	44	3.1	6261	24	ABK31160	Signal transductio
14	43.6	3.1	11084	20	AA522302	Human IL-1ra BAC c
15	43.4	3.1	409	22	AA192030	Human immunoleoti
16	43.4	3.1	13326	24	ABL33712	Human immune syste
17	43.2	3.1	18434	24	ABL34006	Human immune syste
18	43	3.1	201	22	ABA48463	Human breast cell
19	43	3.1	201	22	ABA51474	Human breast cell
20	43	3.1	201	22	ABA66368	Human foetal liver
21	43	3.1	201	22	ABA59502	Human foetal liver
22	43	3.1	201	22	ABA33430	Probe #11896 for g
23	43	3.1	201	22	ABA36424	Probe #14890 for g
24	43	3.1	201	22	AAK14785	Human brain expres
25	43	3.1	201	22	AAK17754	Human brain expres
26	43	3.1	201	22	AAK40526	Human bone marrow
27	43	3.1	201	22	AAK43573	Human bone marrow
28	43	3.1	201	22	AAI21285	Probe #11218 for g
29	43	3.1	201	22	AAI24360	Probe #14293 for g
30	43	3.1	201	22	AAI46560	Probe #15246 used
31	43	3.1	201	22	AAI49630	Probe #18316 used
32	43	3.1	201	22	AAI06989	Probe #5980 used t
33	43	3.1	201	22	AAI09899	Probe #9890 used t
34	43	3.1	201	23	ABS40089	Human liver single
35	43	3.1	201	23	ABS43218	Human liver single
36	43	3.1	201	23	ABS17736	Human genome-deriv
37	43	3.1	395	22	ABA46372	Human breast cell
38	43	3.1	395	22	ABA56938	Human foetal liver
39	43	3.1	395	22	ABA26550	Probe #5016 for ge
40	43	3.1	395	22	AAK05033	Human brain expres
41	43	3.1	395	22	AAK05063	Human bone marrow
42	43	3.1	395	22	AAI15180	Probe #5113 for ge
43	43	3.1	395	22	AAI16508	Probe #5194 used t
44	43	3.1	395	22	AAI04920	Probe #4911 used t
45	43	3.1	395	23	ABS30228	Human liver single

ALIGNMENTS

RESULT 1
AAA99363
ID AAA99363 standard; DNA; 1401 BP.

AC AAA99363;
XX
XX 22-JAN-2001 (first entry)
DT Plant PRAG1 promoter DNA sequence.
DE
DE Plant PRAG1 promoter DNA sequence.
KW Plant promoter; PRAG1; reproductive tissue; transgenic plant; cereal; ds.
XX Pinus radiata.
XX WO200055172-A1.

PD 21-SEP-2000.
XX 17-MAR-2000; 2000WO-NZ00031.
XX 17-MAR-1999; 99NZ-0334715.
XX (CART-) CARTER HOLT HARVEY LTD.
PA (TASM-) TASMAN BIOTECHNOLOGY LTD.
PA (UNMT) UNIV-MICHIGAN TECHNOLOGICAL.
XX Podila GK, Liu J, Karnosky DF;
XX WPI; 2000-594442/56.
XX P-PSDB; AAB26795.

XX WPI; 2000-594442/56.
DR P-PSDB; AAB26795.
XX
PT Novel plant reproductive tissue promoter, useful to produce plants
PT which have a diminished reproductive capacity or which are sterile
XX
PS Claim 2; Page 37-38; 51pp; English.
XX
CC This invention relates to a novel plant promoter gene. The promoter is
CC located in plant reproductive tissue, and the invention includes
CC transgenic plants containing the promoter. The promoter can be used to
CC produce plants which have a diminished reproductive capacity or which are
CC sterile. The constructs can also be used to transform agronomically
CC important plants in which modulation of reproductive capacity
CC (particularly the timing and abundance of flowering) is desirable,
CC e.g. cereals, rice, maize, wheat, barley, oats, rye, soybean and canola.
CC The present sequence represents DNA encoding the plant reproductive
CC promoter (P-AG1) of the invention.
XX
SQ Sequence 1320 BP; 390 A; 269 C; 280 G; 381 T; 0 other;

Query Match 94.2%; Score 1320; DB 21; Length 1320;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1320; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAACCTCGACAGCAAAATATGATTTAGATTATGACTAGAAATAGCATAGATTAAAGCAT 60
DB 1 AAACCTCGACAGCAAAATATGATTTAGATTATGACTAGAAATAGCATAGATTAAAGCAT 60

QY 61 ATACATAAAGCGGTGATATCTCTGACGACCTGCTACTTGAGGAAAGGTAGTGACT 120
DB 61 ATACATAAAGCGGTGATATCTCTGACGACCTGCTACTTGAGGAAAGGTAGTGACT 120

QY 121 CTGCTCAGGTACATTTAGTTGGTAAGTTGGCTTCTGGGTAATAGAGAGTAAA 180
DB 121 CTGCTCAGGTACATTTAGTTGGTAAGTTGGCTTCTGGGTAATAGAGAGTAAA 180

QY 181 GAAGTAAAGGTATTTGACTCTAGTCAAGTACATTCGATTTGCGGGCTTGGAT 240
DB 181 GAAGTAAAGGTATTTGACTCTAGTCAAGTACATTCGATTTGCGGGCTTGGAT 240

QY 241 GGCTTGGGTCGTGTCGAGAGCAACAAATTTAAGAAATATATAAATAAATAAATAA 300
DB 241 GGCTTGGGTCGTGTCGAGAGCAACAAATTTAAGAAATATATAAATAAATAAATAA 300

QY 301 AAATTTAAGTGTGGAGTGAAGCGTGGGCGAGAAATATACACAGAGAGTACTTTAA 360
DB 301 AAATTTAAGTGTGGAGTGAAGCGTGGGCGAGAAATATACACAGAGAGTACTTTAA 360

QY 361 CAATGCGCAACCAAGCAGATTACAACTTGATTTCTGGACCTCGAATACGAGATAATGG 420
DB 361 CAATGCGCAACCAAGCAGATTACAACTTGATTTCTGGACCTCGAATACGAGATAATGG 420

QY 421 TGGTAAGAAATAAGGAAAGTGGAGTGCATTTGAAATGAATGGAGAGCGCAAAATG 480
DB 421 TGGTAAGAAATAAGGAAAGTGGAGTGCATTTGAAATGAATGGAGAGCGCAAAATG 480

QY 481 GAGGACGAATAATGAATTAATGCAAGAGTGCATTTCCCTATTATTTCCAGAAATGA 540
DB 481 GAGGACGAATAATGAATTAATGCAAGAGTGCATTTCCCTATTATTTCCAGAAATGA 540

QY 541 TATGTGGGTCGGCATTCACATGGCGTCCGATTCAGGGGCTGTCATAGCGTCCCTTGA 600
DB 541 TATGTGGGTCGGCATTCACATGGCGTCCGATTCAGGGGCTGTCATAGCGTCCCTTGA 600

QY 601 TTGCAGTGTGGAGTTCGACATGTAACCAATTCATCCCAAAACCTAAATTTA 660
DB 601 TTGCAGTGTGGAGTTCGACATGTAACCAATTCATCCCAAAACCTAAATTTA 660

QY 661 TCCTCTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
DB 661 TCCTCTCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720

QY 721 TCACCTGCTGACACAGTCTTAGTCAATCCATCTGCTTCAAAATAGSCATTATTTTCTTC 780
DB 721 TCACCTGCTGACACAGTCTTAGTCAATCCATCTGCTTCAAAATAGSCATTATTTTCTTC 780

QY 781 TTTCCCTCCGACTGAAAGGCTATCGACCGACCGCTCATCTTCTTCTTCTGCGCAA 840
DB 781 TTTCCCTCCGACTGAAAGGCTATCGACCGACCGCTCATCTTCTTCTTCTGCGCAA 840

QY 841 TTTTCTGCTGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 900
DB 841 TTTTCTGCTGATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 900

QY 901 GGTATCTCTATCTCTCCCTGCAATCGATTGTAGAGAAAGGAGAGGAGGAGGAGATAT 960
DB 901 GGTATCTCTATCTCTCCCTGCAATCGATTGTAGAGAAAGGAGAGGAGGAGGAGATAT 960

QY 961 GTATTGATCAACCTACCCGAAAAAACAATCTGATCAGCCCTGCTCATCTTCTTATAA 1020
DB 961 GTATTGATCAACCTACCCGAAAAAACAATCTGATCAGCCCTGCTCATCTTCTTATAA 1020

QY 1021 TCTCTATCCACTGTTCAATCATTCAGGTTTCTTCCCACTTTTCAAGCAAAAGGCGCCGGA 1080
DB 1021 TCTCTATCCACTGTTCAATCATTCAGGTTTCTTCCCACTTTTCAAGCAAAAGGCGCCGGA 1080

QY 1081 TTGCGCGTGTCTTAGATTTTTCAGTACTTAATTCGACATATTTCCCACTTCCCACTTAA 1140
DB 1081 TTGCGCGTGTCTTAGATTTTTCAGTACTTAATTCGACATATTTCCCACTTCCCACTTAA 1140

QY 1141 TCTCAAAAGATTTGTTGTAGAAACAAAGATTTGATTAATTTTCTTAAAGTTGAGCTTAA 1200
DB 1141 TCTCAAAAGATTTGTTGTAGAAACAAAGATTTGATTAATTTTCTTAAAGTTGAGCTTAA 1200

QY 1201 GGGGTTTGGTACCTAACTTGTGCTTGTGTTATTTTCTTCAAGAACTCGGGCTCGGTCCA 1260
DB 1201 GGGGTTTGGTACCTAACTTGTGCTTGTGTTATTTTCTTCAAGAACTCGGGCTCGGTCCA 1260

QY 1261 ACTGTAGAAAGCAAGCAAGGCTTGCAGCTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 1320
DB 1261 ACTGTAGAAAGCAAGCAAGGCTTGCAGCTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 1320

RESULT 3
ABQ18618/c
ID ABQ18618 standard; DNA; 540 BP.
XX AC ABQ18618;
XX DT 12-JUL-2002 (first entry)
XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 5209.
XX DE Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.
XX OS Homo sapiens.
XX PN WC200218632-A2.
XX PD 07-MAR-2002.
XX PF 01-SEP-2001; 2001WO-EP10074.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PR 05-SEP-2000; 2000DE-1044543.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.
DR

OS Drosophila melanogaster.
XX WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US09231.
XX PR 23-MAR-2000; 2000US-191637P.
XX PR 11-JUL-2000; 2000US-0614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX DR P-PSDB; ABB70899.
XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions -
XX PS Claim 1; SEQ ID NO 39488; 21pp + Sequence Listing; English.
XX CC The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
XX CC sequences (ABU01840-ABU16175) and the encoded proteins
XX CC (ABB57737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 4492 BP; 1158 A; 1245 C; 1141 G; 948 T; 0 other;
Query Match 3.3%; Score 46.4; DB 23; Length 4492;
Best Local Similarity 49.6%; Pred. No. 0.08;
Matches 119; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
QY 804 TCGACCGACCGCTCATCTCTCTCTGCGCAATTTTCTGCTGATCATCATCA 863
Db 2878 TCCAGCGGCGATGTATCAATCTCATCGCCACCAACGATGCCATCGGATCATGTC 2819
QY 864 TTACCATCATCGCCATCCCAACCATCATCATCATGATGATCTCTATCTCTCCCTGGCA 923
Db 2818 TCTTCGTCCTCATCGCCGCTCAGCCTCGTCTCATCATCGTCGCGGCGCATGTGATGCCA 2759
QY 924 ATCGATTGTAGAGAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 983
Db 2758 ACCTGCGGAAGAGCAACGAGAAATGTTATTGAATAATATGTTGCAAACTATAAACAAAGT 2699
QY 984 AACAACTGATCAGCCCTGCTCAATCTTGTCTTATAATCTTATCCACTGTTCAATCAT 1043
Db 2698 TAAAGCAACCAACCAAGCTTTAAATATTAAGATATGATGATGATGATGATGATGATGAT 2639
RESULT 6
ABV50724/c
ID ABV50724 standard; cDNA; 563 BP.
XX AC ABV50724;
XX DT 17-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 50715.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.

XX WO200160860-A2.
XX PN 23-AUG-2001.
XX PD 20-FEB-2001; 2001WO-US05171.
XX PF 17-FEB-2000; 2000US-183319P.
XX PR 16-MAR-2000; 2000US-189862P.
XX PR 25-MAY-2000; 2000US-207454P.
XX PR 09-JUN-2000; 2000US-211314P.
XX PR 18-JUL-2000; 2000US-219007P.
XX PR 13-DEC-2000; 2000US-255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX DR Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer -
XX PS Claim 1; Page 9866; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for:
XX CC (a) assessing whether a patient is afflicted with prostate cancer;
XX CC (b) monitoring the progression of prostate cancer in a patient;
XX CC (c) assessing the efficacy of a test compound to inhibit prostate
XX CC cancer in a patient;
XX CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
XX CC in a patient;
XX CC (e) selecting a composition for inhibiting prostate cancer in a patient;
XX CC (f) assessing the prostate cell carcinogenic potential of a compound;
XX CC (g) determining whether prostate cancer has metastasized in a patient;
XX CC (h) assessing the aggressiveness or indolence of prostate cancer in a
XX CC patient;
XX CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX SQ Sequence 563 BP; 45 A; 107 C; 22 G; 385 T; 4 other;
Query Match 3.3%; Score 45.6; DB 23; Length 563;
Best Local Similarity 49.2%; Pred. No. 0.053;
Matches 120; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 267 AATTTTATAGAAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATA 326
Db 520 AAGAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 461
QY 327 GTGGGGCGAATATATACAGAGAGTACTTTACATGCCGACCAAGGAGGAGGAGGAGGAG 386
Db 460 GACAGGAAAAAGGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 401
QY 387 ACTTCATTTCTGGACCTCGAATACGAGATATGTTGTTAAGATAATAAGGAGGAGGAG 446
Db 400 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 341
QY 447 TGCATTTGAAATGAATCGAGAGCGCAAAATATGAGGAGGAGGAGGAGGAGGAGGAGGAG 506
Db 340 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 281
QY 507 AAGA 510
Db 280 AAAA 277
RESULT 7
ABX22233/c
ID ABX22233 standard; cDNA; 519 BP.
XX

XX PS Claim 1; SEQ ID No 63; 24pp; English.

XX CC The present invention relates to chemically modified DNA sequences of

XX CC signal transduction associated genes. The DNA sequences are chemically

XX CC modified using a solution of bisulphite, hydrogen sulphite or

XX CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers

XX CC for detecting the cytosine methylation state (CpG islands) of these

XX CC genes, and a method for the diagnosis and/or therapy of genetic and

XX CC epigenetic parameters of genes associated with signal transduction.

XX CC The genomic DNA can be obtained from cells or cellular components which

XX CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,

XX CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from

XX CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,

XX CC histologic object slides, and all their possible combinations. The

XX CC sequences of the invention are useful for the diagnosis and therapy of

XX CC diseases associated with signal transduction e.g. solid tumours and

XX CC cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA

XX CC sequences of different genes associated with signal transduction, or

XX CC their complementary sequences.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from the

XX CC European Patent Office.

XX SQ Sequence 15923 BP; 3153 A; 441 C; 4469 G; 7860 T; 0 other;

Query Match 3.2%; Score 44.4; DB 24; Length 15923;

Best Local Similarity 50.8%; Pred. No. 0.47; Indels 3; Gaps 1;

Matches 133; Conservative 0; Mismatches 126;

QY 97 GTACTTGAGAAAGGTAGTCTGCTCAGTCAATAGTTGGTAAAGTTGGCTTGG 156

Db 10769 GAAATGAGTATGTTAGGGGATTTAAGGATTTTATTTGTTTGGAAATTTGTTGCG 10828

QY 157 CTTCCTGGTAAATGAGAGTAAGAGTAAGAGTAAAGGATTTTACTAGTCAAGTACATG 216

Db 10829 TTTATTTAGTACGAGTATGATGCTTAAAGTTTATTTATGTTTGAAGTATTTT 10888

QY 217 GATTCCTTTGTCGGGCTTGGATGCTTGGTTCGTGAGAGCAACAAATTTATAAG 276

Db 10889 ATTTTCTGTTGTTAGTATGATTTATTTTGTGTTTATTTTATTTTATTTAA 10945

QY 277 AATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 336

Db 10946 AATATTAATTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11005

QY 337 AATATACACAGAGTACTTT 358

Db 11006 AAAAAATACGTTGTATGATTTT 11027

RESULT 11

Id AAS46643 standard; DNA; 6261 BP.

XX AAS46643;

XX 18-DEC-2001 (first entry)

XX Tumour suppressor gene derived chemically modified sequence #365.

XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;

XX cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;

XX cytosine methylation; ds.

XX Homo sapiens.

XX WO200168912-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2001WO-EP02955.

PR 15-MAR-2000; 2000DE-1013847.

PR 06-APR-2000; 2000DE-1019058.

PR 07-APR-2000; 2000DE-1019173.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.

XX Olek A, Pipenbrock C, Berlin K;

XX WPI; 2001-602752/68.

XX Fragments of chemically modified genes associated with tumour suppressor

XX genes and oncogenes, useful in designing primers and probes for

XX analysing diseases associated with cytosine methylation state e.g.

XX cancer

PS Claim 1; SEQ ID No 365; 27pp; English.

XX The invention relates to a nucleic acid comprising a sequence of 18

XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with

XX bisulphite, of genes associated with tumour suppression and

XX oncogenes having a sequence taken from 536 (actually 533 since

XX numbers 406, 458 and 500 are missing from the sequence listing) sequences

XX (SS) and sequences complementary to (SS). The nucleic acid may be a

XX peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may

XX form part of a set of probes for detecting the cytosine methylation state

XX and/or single nucleotide polymorphisms and also to be used in an

XX array for analysing diseases associated with CpG dinucleotides e.g.

XX cancers and tumours. The probes can also be used in a method for

XX ascertaining genetic and/or epigenetic parameters for the diagnosis

XX and/or therapy of existing diseases or the predisposition to specific

XX diseases, by analysing cytosine methylations. The parameters may be

XX compared to another set of genetic and/or epigenetic parameters, the

XX differences serving as basis for diagnosis and/or prognosis events which

XX are disadvantageous to patients. The present sequence is one of the

XX 533 genomic sequences derived from tumour suppressor genes and

XX oncogenes.

XX Note: The sequence data for this patent did not form part

XX of the printed specification, but was obtained in electronic

XX format directly from WIPO at

XX ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 6261 BP; 1716 A; 190 C; 1696 G; 2659 T; 0 other;

Query Match 3.1%; Score 44; DB 22; Length 6261;

Best Local Similarity 50.5%; Pred. No. 0.4; Indels 0; Gaps 0;

Matches 107; Conservative 0; Mismatches 105;

QY 146 GGTTCGCTTGGCTTCTCGGTAATATGAGAGTAAGAGTAAGAGTAAGAGTAAGAGT 205

Db 1216 GGTCTTTTGTAGTTTATGTTTATTCGGAGGTTGAGGTAGGAGATGTTGTAATTCGGGA 1275

QY 206 CAAGTACATTTGATTCCTTTTTCGGGCTTGGATCGGCTTGGTTCGTGAGAGCCAA 265

Db 1276 GCGCGAGTTTGTAGTTCGAGTCGAGATTCGCGTTATTTATTTAGTTTGGGGATAGAGCA 1335

QY 266 CAATTTATAAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 325

Db 1336 GATTTTATTTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1395

QY 326 GGTGGGCGCAAAATATACAGAGAGTACTT 357

Db 1396 AGGTAAGTAATTTTATATAACGAAGAATATT 1427

RESULT 12

Id AAS61046 standard; DNA; 6261 BP.

XX AAS61046;

XX 29-JAN-2002 (first entry)

CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 409 BP; 183 A; 50 C; 77 G; 97 T; 2 other;
Query Match 3.1%; Score 43.4; DB 22; Length 409;
Best Local Similarity 52.9%; Pred. No. 0.18; 82; Indels 0; Gaps 0;
Matches 92; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 151 GCTTGGCTTCTGGGTAATATGAGAGTAAGAAAGTAAAGGTATTTGACTCTAGTCAAGT 210
Db 164 GTTTGGGTTTGGGTTAGAGGAGATGAATAAGAGAGATTGTATTTAACTAGCTAGG 223
QY 211 ACATTGGATTGCCCTTTGTCCGGGCTTGGATGGCTTGGTTCGTGTGAGAGCCACAAATT 270
Db 224 AAAATGGTGCACTGAATAAAGTTTACTGCAGGGGTTTTCCCCCATGACAAAGCANAATAAA 283
QY 271 TATAGAAATATATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 324
Db 284 AA 337
Search completed: January 21, 2004, 17:31:47
Job time : 362.266 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 21, 2004, 17:02:36 ; Search time 2794.45 Seconds
(without alignments)
12185.050 Million cell updates/sec

Title: US-09-936-869-2
Perfect score: 1401
Sequence: 1 aaactgacagcaaatatga.....ggaaattttgtacaaatc 1401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpi.*
- 7: em_estro.*
- 8: em_hic.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_hic.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_phg.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	166	11.8	526	9 AW754659	AW754659 PC04G10 P
2	61.4	4.4	1101	29 CNS0039G	AL063921 Drosophil
3	57.2	4.1	987	29 CNS00418	AL06537 Drosophil
4	56.2	4.0	1201	13 BX461310	BX461310 BX461310

5	55.2	3.9	1101	29 CNS003B2	AL064080 Drosophil
6	55.2	3.9	1101	29 CNS00DG1	AL068971 Drosophil
7	54.8	3.9	1041	9 AL559034	AL559034 Drosophil
8	54	3.9	964	29 CNS006N9	AL065781 Drosophil
9	54	3.9	1101	29 CNS0174K	AL107582 Drosophil
10	54	3.9	1101	29 CNS017KE	AL108152 Drosophil
11	53.4	3.8	1101	29 CNS0100X	AL098379 Drosophil
12	52.8	3.8	997	29 CNS000TE	AL060767 Drosophil
13	52.6	3.8	634	29 BX122696	BX122696 Drosophil
14	52.6	3.8	1101	29 CNS0182P	BX122696 Drosophil
15	52.6	3.8	1114	13 BX407949	AL108811 Drosophil
16	52.6	3.8	1201	9 AL566565	AL566565 Drosophil
17	52	3.7	783	29 CNS000AIS	AL055833 Drosophil
18	52	3.7	797	13 BX437575	BX437575 Drosophil
19	52	3.7	1101	29 CNS0181N	AL108773 Drosophil
20	51.8	3.7	623	28 AQ157491	AQ157491 nxbx0009H
21	51.8	3.7	830	29 CNS06P72	AL409029 T7 end of
22	51.8	3.7	1099	13 BX456575	BX456575 Drosophil
23	51.6	3.7	946	13 BX416277	BX416277 Drosophil
24	51.6	3.7	1201	13 BX461128	BX461128 Drosophil
25	51.4	3.7	1080	28 AQ738730	AQ738730 HS_5382_B
26	51.4	3.7	1101	29 CNS0039R	AL063932 Drosophil
27	51.2	3.7	997	29 CNS000TE	AL060767 Drosophil
28	51.2	3.7	1204	29 CNS016E2	AL108628 Drosophil
29	51	3.6	673	29 CNS03XD5	AL264866 Tetradon
30	50.8	3.6	1011	13 BX349844	BX349844 Drosophil
31	50.6	3.6	929	13 BX408907	BX408907 Drosophil
32	50.6	3.6	938	29 CNS006TJ	AL065906 Drosophil
33	50.4	3.6	626	28 AZ020346	AZ020346 RPI-23-3
34	50.4	3.6	924	29 CNS01GHN	AL143148 Anopheles
35	50.4	3.6	1001	29 CNS0064G	AL063781 Drosophil
36	50.2	3.6	1101	29 CNS017KX	AL108171 Drosophil
37	50	3.6	1049	13 BX353130	BX353130 Drosophil
38	50	3.6	1101	29 CNS0106X	AL098595 Drosophil
39	50	3.6	1201	13 BX335607	BX335607 Drosophil
40	50	3.6	1201	13 BX385867	BX385867 Drosophil
41	49.6	3.5	658	28 AZ851742	AZ851742 2M0154E07
42	49.6	3.5	735	29 AG095704	AG095704 Pan trogl
43	49.6	3.5	1201	13 BX460090	BX460090 Drosophil
44	49.2	3.5	732	28 BZ146517	BZ146517 CH230-453
45	49.2	3.5	760	29 AG056350	AG056350 Pan trogl

ALIGNMENTS

RESULT 1
AW754659

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AW754659 526 bp mRNA linear EST 01-MAY-2000
PC04G10 Pine Triplex pollen cone library Pinus taeda cDNA clone
PC04G10, mRNA sequence.

AW754659.1 GI:7676379

EST.

Pinus taeda (loblolly pine)

Pinus taeda

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

1 (bases 1 to 526)

Whetten,R.W., Kinlaw,C.S., Retzel,E. and Sederoff,R.R.

The Pine Gene Discovery Project

Unpublished

Contact: Ross Whetten

Forest Biotechnology Group

North Carolina State University

Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh

,NC, 27695-8008

Tel: 919-515-7800

Fax: 919-515-7801

Email: rosswhetten@unity.ncsu.edu

Seq primer: 5' lambda Triplex2 Sequencing Primer.

Location/Qualifiers

1..526

FEATURES
Source

```

/organism="Pinus taeda"
/mol_type="mRNA"
/db_xref="taxon:3352"
/clone="PC04G10"
/dev_stage="immature"
/lab_host="E. coli BM25.8"
/clone_lib="Pine triplex pollen cone library"
/note="Organ: pollen cone; Vector: Lambda triplex; Site_1:
SfiI (A); Site_2: SfiI (B); Immature pollen cones were
collected in the early spring, frozen and used for mRNA
isolation. The SMART-PCR method (Clontech) was used to
prepare a library from 1 ug total RNA, using the Lambda
Triplex vector. Plasmid subclones in pTriplex were
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."
BASE COUNT      134 a    99 c    128 g    146 t    19 others
ORIGIN
Query Match      11.8%; Score 166; DB 9; Length 526;
Best Local Similarity 97.1%; Pred. No. 3.1e-17;
Matches 169; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1228 GTATTGTTCTCAGAACTCGGCTGTCCTCACTGTAGGAACGACCAAGGGG 1287
DB 154 GTCTTAGAATTTCAAGACTCGGCTGCTCCAACTGTAGGAACGACCAAGGGG 213
QY 1288 TTGCAGCTTTGCTGTGCTGTGCGCCCACTGCTTTTGGACTGCTATTAGTTCAG 1347
DB 214 TTGCAGCTTTGCTGTGCTGTGCGCCCACTGCTTTTGGACTGCTATTAGTTCAG 273
QY 1348 CTTGTTTTCATGCTGTGAGATCTGCGCGGAATTTGTGACAAATC 1401
DB 274 CTTGTTTTCATGCTGTGAGATTTGTGCGCGGAATTTGTGACAAATC 327

CNS0039G      1101 bp      DNA      linear      GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BAC08K10 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL063921
AL063921.1      GI:4941778
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osagawa and
Aaron Mammocser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
location/Qualifiers
1..1101
/organism="Drosophila melanogaster"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08K10"
/clone_lib="RPCI-98"
/note="end : TET3"
BASE COUNT      201 a    64 c    131 g    202 t    503 others
ORIGIN
Query Match      4.4%; Score 61.4; DB 29; Length 1101;
Best Local Similarity 17.9%; Pred. No. 2;
Matches 140; Conservative 295; Mismatches 339; Indels 9; Gaps 1;
QY 17 ATGATTTAGATTATGACCTAGAAATAAGCATAGATTAAGCATATACATACAAAGCGGT 76
DB 1013 TTTWMTAWRAGRDGAGRKRDAAATDAGAGRRGRRKXKDKDKDGGDDDKGGKK 954
QY 77 GATATACCTCTGACCTGCCACTGTACTTGGAGAAAGTAGTGGACTCTCTCAGGTACATTA 136
DB 953 KKAACAATKATKWDWDKDKWDGAKDKADDDDDGAGDKDDDKGDKADDDDDTDTGDK 894
QY 137 GTTTGGTAAGTTGGCTTCTGCGGTAATATGAGAAGTAAGAAGTAAGAAGTATTT 196
DB 893 DDDDKKDDDDKAGTGGDATWAAATDWWWWGADADWWTWDAADDDWADDDWADAW 834
QY 197 GACTCTAGTCAAGTACATTTGCTTTGCGGGCTTGGATGGCTTGGGTTCGTGTG 256
DB 833 KWDDAWANGARTADRRDWDGAGRKGRKRDADDDDDAADDRAATTTT 774
QY 257 AGAGCCCAATTTATAGAAATATATAAATAAATAAATAAATAAATAAATAAATAA 316
DB 773 TRDIDDKWKYDTTWRWAADRTWDRDDDDDRAGTAGRKWRRTWKRWKRRDTRWDDAD 714
QY 317 AGTGAACG-----GTGGGGCAGAAATATACACAGAGAGTACTTTAACAATCGG 367
DB 713 ADDTARDRRRRGGDAGAGKKTGRKRRRRDRATWDRDADWAAWMTTDTTDDW 654
QY 368 CAACAGGCGAGATTCACAACTTGATTTCTGGACCTCGATACGATAGATAGTGTGTAAG 427
DB 653 DKRRRRKGRARRRTTARAANDWTTWAKMDWAKMDWKTTRADRWDRWAADTTIDARKAD 594
QY 428 AATAAAGGAGAGTGGAGTGCATTTGAAAATGAATGAGAGCGCACAAATCGAGACG 487
DB 593 DWAKARAWRARRDRARAARADRRWTTKGKTTTATTTTAAWAAWAAWAAWAAWATTT 534
QY 488 AATAAATGAATATAATGCAAGAGTGCATTTCCCTATTATTTCCAGAAATGTATATG 547
DB 533 WTTTWTTTTWTTTTWTAAWAAWATTAWAAWAAWAAWAAWAAWATTTT 474
QY 548 GTTCGGCATTCACATGGGCGTGCATTCAGGGGGTGTCAATAGCGGCTCTTTGATCGAT 607
DB 473 TTTWAAWAAWATTAWTTTWTTTTWTAAWATTTTWTAAWATTTTWTAAWATTTT 414
QY 608 GTGGGAGTTGCAACATGTACCAACAAATCCATTCATCCCAAAACCTAAATTTATCCTC 667
DB 413 AAAAAAWWWTTTATATKCCCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 354
QY 668 CATTACTATTACCTACACCTATACCTAGTAGTAATAATATGTCCTGCTGTAACTCCAC 727
DB 353 YYYYYYYYVCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 294
QY 728 CCTGCACAGCTCTTAGTCAATCATCTGCTTCAATATAGCATTTATTTCTTCTTCTT 787
DB 293 CYTTTTCNCNCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 234
QY 788 TCC 790
DB 233 YCY 231

```

```

RESULT 3
CNS00418
LOCUS      Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION

```

BACR09C16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
 AL066537
 AL066537.1 GI:4942778
 GSS.
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 987)
 Genoscope.
 Direct Submission
 Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segre@genoscope.cns.fr
 - web : www.genoscope.cns.fr)
 Determination of this BAC-end sequence was carried out as part of a
 collaboration with the Berkeley Drosophila Genome Project (BDGP).
 The BDGP is constructing a physical map of the Drosophila
 melanogaster genome using these BACs. For further information
 please see http://www.fruitfly.org The BDGP Drosophila
 melanogaster BAC library was prepared by Kazuo Osoegawa and
 Aaron Mammeter in Pieter de Jong's laboratory in the Department of
 Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
 NY. The library is named RPCI-98 and was constructed by partial
 EcoRI digestion of Drosophila DNA provided by the BDGP from the
 isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
 P1 and EST libraries. A more detailed description of the library
 and how to order individual BAC clones, the entire library, or
 filters for hybridization from the BACPAC Resource Center can be
 found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers	1..987
source	/organism="Drosophila melanogaster"
	/mol_type="Genomic DNA"
	/db_xref="taxon:7227"
	/clone="BACR09C16"
	/clone_lib="RPCI-98"
	/note="end : TET3"

BASE COUNT 238 a 162 c 17 g 177 t 393 others

ORIGIN

```

Query Match      4.1%; Score 57.2; DB 29; Length 987;
Best Local Similarity 27.1%; Pred. No. 9.8;
Matches 180; Conservative 94; Mismatches 386; Indels 5; Gaps 1;
```

Qy 258 GARCCACACTTTATAGRAATATATAAATAAAAAATAAAAAATTAAAGTGTGGAA 317
 |||||:::||||:
 Db 52 GA AAAA VAAAAANNDWCGANNNNNNAANAANAANNNAATATATANA AAAAAADA 111
 |||||:::||||:
 Qy 318 GTGAAAACGGTGGGCAGAAATATACACAGAAGAGTACTTTAACAAATGCCCAACAAGGC 377
 |||||:::||||:
 Db 112 ATA NAAAAA AAAAAAAAAAATAACAAAYAAASHGNMNNNGAGGGAANNAANTNN 171
 |||||:::||||:
 Qy 378 AGATTCAACTTGATTTCCTGGACCTCGAATACGAGTAATGGTGGTAGAATAATAAGGA 437
 |||||:::||||:
 Db 172 NNNAANNCTTNVANNGAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 231
 |||||:::||||:
 Qy 438 AGAGTGGAGTGCTATTGAAATGAATGGAGAGCGCACAAAATGGAGGACGAATAATGAA 497
 |||||:::||||:
 Db 232 AAAAAA NAAAAANTNRAGANANANANNTNATATAATAGNRAAANA AAAAAA AAAAA 291
 |||||:::||||:
 Qy 498 ATATAATCAAGAGTGCATTTCCTCATTTATTTCCAGAAATGTATATGTGGGGTGGCATT 557
 |||||:::||||:
 Db 292 AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAAA AAAAA 351
 |||||:::||||:
 Qy 558 CACATGGCGGTGCATTTCAGGGGGTGTCATACGGCTCTTGTGATTGCGAGTGGGAGTTG 617
 |||||:::||||:
 Db 352 NANNNNNNAANNVANCCAGANNNAANNNNNNNNAANNNNNNAANNNNAAGANANNNA 411
 |||||:::||||:
 Qy 618 CAACATGTACCAACAAATCCATTATCCCAAACCTAAATTTATTCCTCTCCATTACTATT 677
 |||||:::||||:
 Db 412 NANNNAANNAANGANNANGNATCYTCYCCYTCYTCYTCYTCYTCYTCYTCYTCYTCYTC 471
 |||||:::||||:

Qy	678	ACCTACACCTATACCTAGTAAATGATGTCTTCGGCCTTGTAACCTCCCACTGCCTGCACCG	737
Dd	472	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	531
Qy	738	TCTTAGTCAATCATCTGCCTTCAAATAGGCATT-----ATTTTGTTCTTTCCCTCCGA	792
Dd	532	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	591
Qy	793	CYCYVCCCCCCCVCYCYCYCYCYCTCCCTTCCTTCCTTCCTGGCAATTTTTTCTGCTG	852
Dd	592	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	651
Qy	853	GATCAATCATTTACCATCATTCGCCATCCGCCACATCATCATCATGATGATGATCTCTATC	912
Dd	652	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	912
Qy	913	TYTCC 917	711
Dd	712	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	711
		1201 bp mRNA linear EST 22-MAY-2003	
RESULT 4	BX461310	Homo sapiens FETAL BRAIN Homo sapiens cDNA clone	
LOCUS	CSODP025YA16	5-PRIME, mRNA sequence.	
DEFINITION	AX461310		
ACCESSION	AX461310.1	GI:31031291	
VERSION	EST.		
KEYWORDS	Homo sapiens		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Li,W.B., Gruber,C.; Jesses,J. and Polayes,D.		
TITLE	1 (bases 1 to 1201)		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished		
FEATURES	Contact: Genoscope		
source	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr		
	Library was constructed by Life Technologies, a division of		
	Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :		
	http://fulllength.invitrogen.com/ Invitrogen Corporation 1600		
	Faraday Avenue Genoscope sequence ID : CSODF025BA08QP1.		
	Location/Qualifiers		
	1..1201		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="CSODF025YA16"		
	/tissue types="FETAL BRAIN"		
	/dev stage="fetal"		
	/clone lib="Homo sapiens FETAL BRAIN"		
	/note="Organ: brain; Vector: pCMVSORT 6; 1st strand cDNA		
	was primed with a NotI-oligo(dT) primer. Five prime end		
	enriched, double-strand cDNA was digested with Not I and		
	cloned into the Not I and EcoRV sites of the pCMVSORT 6		
	vector. Library was not normalized."		
BASE COUNT	472 a 108 c 138 g 167 t 316 others		
ORIGIN			
Query Match	4.0%; Score 56.2; DB 13; Length 1201;		
Best Local Similarity	21.6%; Pred. No. 13;		
Matches	81; Conservative 139; Mismatches 155; Indels 0; Gaps 0		
Qy	141	GGAAGGTTGGCTTCTCGGGTAATATGAGAAGTAAGATAAAGTAAAGTATTGACT	200
Dd	792	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	841
Qy	201	CTAGTCAAGTACATGGATTGCCITTGTGCGGGCTTGATGCTGTTGGTTCGTGTGAGAA	260

[illegible]

RESULT 9
CNS0174K 1101 bp DNA linear GSS 26-JUL-1999
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN17P20 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL107582
VERSION AL107582.1 GI:5627886
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
REFERENCE

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 1101)	Genoscope.		
	Direct Submission			
	Submitted (23-Jul-1999)	Genoscope - Centre National de Sequencage :		
	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr			
	- Web : www.genoscope.cns.fr)			
	Determination of this BAC-end sequence was carried out as part of a			
	collaboration with the European Drosophila Genome Project (EDGP) -			
	http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC			
	library (Dros BAC) was made by Alain Billaud at C3PH (Centre			
	d'Etude du Polymorphisme Humain) with funding provided by a MRC			
	project grant. The DNA was prepared from embryos by Alain Sucheton			
	and Genevieve Payan. It has been constructed in the vector			
	pBeloBAC11.			

```

FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN17P20"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : SP6"

```

BASE COUNT	365 a	90 c	235 g	206 t	205 others
ORIGIN					
Query-Match		3.9%	score 54;	DB 29;	Length 1101;
Best Local Similarity		42.5%;	Pred. No. 31;		

253 TGTGAGAGCCACAAATTATAGAAATATATATAAATAAAAAATAAAAATTTAAGTGT 312
 216 TATATAGCATTAATAATTAGGGWTAATAAAAAAAWATAAGCAATTAATAAARGAA 275

QY 313 TCGAAGTGA AACCGTGGGCGAAATATACAGAAAGTAGTCTTTTACAAATGGCGAAC 372

DB 276 AATTAATAAATAAAGAGCAAAATATTAATAAAGAGTARWTATTTWAAAAATATGGAAT 315

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 16:52:57 ; Search time 94.7682 Seconds
(without alignments)
6525.161 Million cell updates/sec

Title: US-09-936-869-2
Perfect score: 1401
Sequence: 1 aaactgcacgaatgatga.....ggaaattttgtgtacaaatc 1401

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/2/ina/6CTUS COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	73.4	5.2	7218	1	US-08-232-463-14
C 2	61	4.4	7218	1	US-08-232-463-14
C 3	43.6	3.1	176373	3	US-09-128-155-17
C 4	42	3.0	152331	3	US-09-128-155-16
C 5	41	2.9	2255	2	US-08-741-134-1
C 6	40	2.9	87350	3	US-08-781-891-79
C 7	40	2.9	87350	4	US-09-618-166-79
C 8	40	2.9	87543	4	US-09-731-211-3
C 9	39.4	2.8	435	4	US-09-397-787-171
C 10	39.4	2.8	1024	4	US-09-328-475C-64
C 11	39.2	2.8	1753	3	US-09-058-260-21
C 12	39.2	2.8	1756	3	US-09-058-260-31
C 13	39.2	2.8	1776	3	US-09-058-260-23
C 14	39.2	2.8	1952	3	US-09-058-260-5
C 15	39.2	2.8	4090	2	US-08-781-802-5
C 16	39.2	2.8	4090	3	US-08-694-078-5
C 17	38.8	2.8	3564	2	US-07-594-321C-9
C 18	38.8	2.8	3564	4	US-09-197-948-9
C 19	38.6	2.8	202001	4	US-09-734-674-3
C 20	38.2	2.7	289	3	US-09-007-005-17
C 21	38.2	2.7	289	3	US-09-244-796-17
C 22	37.6	2.7	1896	3	US-09-058-260-3
C 23	37.6	2.7	6263	2	US-08-781-802-3
C 24	37.6	2.7	6263	3	US-08-694-078-3
C 25	37.4	2.7	475	4	US-09-702-705-36
C 26	37.4	2.7	475	4	US-09-736-457-36
C 27	37.4	2.7	4235	4	US-09-702-705-317

C 28	37.4	2.7	4235	4	US-09-736-457-317
C 29	37.4	2.7	72604	4	US-09-268-992-7
C 30	37.4	2.7	72604	4	US-09-657-474-7
C 31	37	2.6	5720	4	US-09-442-100-1
C 32	37	2.6	5720	4	US-08-939-106-1
C 33	36.8	2.6	289	3	US-09-007-005-17
C 34	36.8	2.6	289	3	US-09-244-796-17
C 35	36.8	2.6	246240	2	US-08-724-394A-20
C 36	36.8	2.6	246240	2	US-08-724-394A-21
C 37	36.8	2.6	246240	2	US-08-724-394A-22
C 38	36.6	2.6	1664976	4	US-08-916-421B-1
C 39	36.4	2.6	658	3	US-08-998-416-595
C 40	36.4	2.6	1164	4	US-09-538-871-3
C 41	36.4	2.6	1232	4	US-09-538-871-1
C 42	36.2	2.6	2013	4	US-09-134-001C-103
C 43	36.2	2.6	176373	3	US-09-128-155-17
C 44	36	2.6	2518	3	US-09-433-699-3
C 45	36	2.6	1230025	4	US-09-198-452A-1

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F18
US-08-232-463-14

Query Match 5.2%; Score 73.4; DB 1; Length 7218;

Best Local Similarity 6.2%; Pred. No. 8.5e-11;
Matches 26; Conservative 236; Mismatches 157; Indels 0; Gaps 0;
QY 93 CACTGTACTTGAGGAAGGTAGTGGACTCTGCTCAGGTACATAGTTGGTAAAGTTGGC 152
Db 1487 CACTGTAAATACCTACTCTATGCAAGTAGTAAAGAGATAGAAGAAATTTGGTACRRRRR 1428
QY 153 TTGGCTTCTGGTAATATAGAGAGTAAAGAGTAAAGGTAATTTGACTCTAGTCAAGTAC 212
Db 1427 RRR 1368
QY 213 ATTGGATGCTTTGCTGGGGCTTGGATGCTTGGCTTGGCTGCTGAGAACCAATTTA 272
Db 1367 RRR 1308
QY 273 TAAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 332
Db 1307 RRR 1248
QY 333 CAGAAATATACAGAGAGTACTTTTAAACAATGCGCAACCAAGGAGATTCACAATTGA 392
Db 1247 RRR 1188
QY 393 TTCTGCACTCGAATACGAGATAATGGTGAAGAAATAAGGAGAGTGGAGTGCAAT 452
Db 1187 RRR 1128
QY 453 TGAATATGAGGAGCGCACAAATGGAGGACGAATAAATGAATAATAATGAACAAG 511
Db 1127 RRR 1069
RESULT 2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:

; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14
Query Match 4.4%; Score 61; DB 1; Length 7218;
Best Local Similarity 2.2%; Pred. No. 3e-07;
Matches 7; Conservative 200; Mismatches 110; Indels 0; Gaps 0;
QY 603 GCAGTGTGGAGTGGCAACATGATACCAACAAATCCATTCATCCAAACCTAAATTTATC 662
Db 1055 GGAGCTTGGCATTT 1114
QY 663 CTCTCCATTACTATTACTACACCTATACCTAGTAAATATGTCCTGCTTGTAACTCTC 722
Db 1115 YY 1174
QY 723 CACTGCTGCACAGCTTTAGTCAATCAATCGCTTCAATAGGCAATTTTGTCTT 782
Db 1175 YY 1234
QY 783 TCCCTCCGACTGAAGGCTATCGACGACGCGCTCATCTTCTTCTTCTGCGCAAT 842
Db 1235 YY 1294
QY 843 TTTTCTGCTGATCATCATCATTTACATTCATCATCGCCATCCCAACCATCATCATGATG 902
Db 1295 YY 1354
QY 903 TATCTCTATCTCTCCCT 919
Db 1355 YYYYYYYYYYYYYYYYYY 1371
RESULT 3
US-09-128-155-17/c
; Sequence 17, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Fan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17
Query Match 3.1%; Score 43.6; DB 3; Length 176373;
Best Local Similarity 48.8%; Pred. No. 0.15;
Matches 118; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 274 AGAATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 333
Db 122513 AAGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 122454
QY 334 AGAAATATACAGAGAGAGTACTTTTAAACATGCGCAACCAAGGAGATTCAACTTGAT 393

RESULT 5
US-08-741-134-1/c
; Sequence 1, Application US/08741134
; Patent No. 5861498
; GENERAL INFORMATION:

```

RESULT 6
US-08-781-891-79/c
; Sequence 79, Application US/08781891
; Patent No. 6090620
;
GENERAL INFORMATION:
;
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Yu, Chang-En
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
;
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
;
TITLE OF INVENTION: WERNER'S SYNDROME
;
NUMBER OF SEQUENCES: 209
;
CORRESPONDENCE ADDRESS:
;

```



```
NAME/KEY: unsure
LOCATION: 12742
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 53384
OTHER INFORMATION: unknown
LOCATION: 29370
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29422
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29979
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29980
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 29981
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30136
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 30140
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31205
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31206
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 31160
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34066
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 34072
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 36816
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 39020
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 42164
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 42459
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46808
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46823
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 46826
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 47291
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 52786
OTHER INFORMATION: unknown
NAME/KEY: unsure

LOCATION: 52787
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 53384
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 54684
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59215
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59235
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 59242
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 63290
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 66614
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68660
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68697
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68718
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68733
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 68739
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 69785
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79114
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 79198
OTHER INFORMATION: unknown
NAME/KEY: unsure
LOCATION: 86336
OTHER INFORMATION: unknown
OTHER INFORMATION:
US-09-791-211-3

Query Match      2.9%; Score 40; DB 4; Length 87543;
Best Local Similarity 50.8%; Pred. No. 1.1;
Matches 120; Conservative 0; Mismatches 115; Indels 1; Gaps 1;

QY 269 TTTATAGAAATATATAAAATATAAAATATAAAATTTTAAAGTGTGGAGTGAAACGGT 328
DB 49995 TTGGTAAATATAAATATATATATAATAAAAGAAATTTTAAATCTTAAAAACGT 49936

QY 329 GGGGAGAAATATACACAGAGAGTACTTTAA-CAATGCGCAACCAAGGAGATTCAAA 387
DB 49935 TTTAAATAATTTTAAATAAGTGAATACAGAAATATAATAATAAA 49876

QY 388 CTTGATTTCTGGACCTCGAATACAGATATATGGTGTGAAGAAATAAAGGAGAGTGGAGT 447
DB 49875 TTATGTTAAAAAGAGAAATAGAGGTGAAAGAGAGAGAAATTTTAAAGAGAAATGAAG 49816

QY 448 GCATTTTGAATGAATGGAGAGCGCACAAAATGGAGGACGAATAAATGAATATAA 503
DB 49815 AATAAAGAAATGATGATAAAGTGAATAATTTGAAGCTGGGTATATGAAGATCCAA 49760
```

RESULT 9
US-09-397-787-171/c
; Sequence 171, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 171
; LENGTH: 435
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-171
Query Match 2.8%; Score 39.4; DB 4; Length 435;
Best Local Similarity 71.2%; Pred. No. 0.11; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 21
QY 825 TTCTTCTCTGGCAATTTTCTGCTGGATCATCATCATATACCATCGCCATCCCA 884
DB 372 TTCTCATCTGGCGCTTTTCTTCAGCTTCCTCATCATCAAAATCATCATCATCATCT 313
QY 885 CCATCATCATCAT 897
DB 312 TCATCATCATCT 300
RESULT 10
US-09-328-475C-64/c
; Sequence 64, Application US/09328475C
; Patent No. 6476207
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jimmy
; APPLICANT: Astel, Jon H.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna M.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Steinmann, Kathleen E.
; TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
; TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
; FILE REFERENCE: 1532.002/200130.463
; CURRENT APPLICATION NUMBER: US/09/328,475C
; CURRENT FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo Sapien
; NAME/KEY: misc feature
; LOCATION: (1)...(1024)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-64
Query Match 2.8%; Score 39.4; DB 4; Length 1024;
Best Local Similarity 71.2%; Pred. No. 0.17; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 21
QY 825 TTCTTCTCTGGCAATTTTCTGCTGGATCATCATCATATACCATCGCCATCCCA 884
DB 440 TTCTCATCTGGCGCTTTTCTTCAGCTTCCTCATCATCAAAATCATCATCATCATCT 381

QY 885 CCATCATCATCAT 897
DB 380 TCATCATCATCT 368
RESULT 11
US-09-058-260-21
; Sequence 21, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fonstein, Michael
; APPLICANT: Vonstein, Veronika
; APPLICANT: Demitrian, David
; APPLICANT: Casadaban, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; FILE REFERENCE: 95-963-H
; CURRENT APPLICATION NUMBER: US/09/058,260B
; CURRENT FILING DATE: 1999-04-10
; EARLIER APPLICATION NUMBER: 60/001,995
; EARLIER FILING DATE: 1996-08-07
; EARLIER APPLICATION NUMBER: 60/009,704
; EARLIER FILING DATE: 1996-01-11
; EARLIER APPLICATION NUMBER: 60/019,580
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: 08/694,078
; EARLIER FILING DATE: 1996-08-08
; EARLIER APPLICATION NUMBER: 08/781,802
; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1753
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: cloned esterase
; OTHER INFORMATION: gene from bacteria E013
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)...(1630)
US-09-058-260-21
Query Match 2.8%; Score 39.2; DB 3; Length 1753;
Best Local Similarity 56.1%; Pred. No. 0.25; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 58
QY 205 TCAAGTACATTCGATTCCTTTTCGGGCTTGGATGGCTTGGTTCGTGAGAGCCA 264
DB 24 TCGGTGCGATGGATTCCTTCAGGGGAACCTTTTAAACACTTGAGTTTGACACCACTCT 83
QY 265 ACAATTTATAAGAAATATATAATAATAATAATAATAATAATAATAATAATAATA 324
DB 84 TAATCATTTAAGATTTAAATGAAATTAATAATAATAATAATAATAATAATAATAATA 143
QY 325 CGGTGGGCGAGA 336
DB 144 CGTGTGTGAAA 155
RESULT 12
US-09-058-260-31
; Sequence 31, Application US/09058260B
; Patent No. 6218167
; GENERAL INFORMATION:
; APPLICANT: Allen, Larry
; APPLICANT: Aikens, John
; APPLICANT: Fonstein, Michael
; APPLICANT: Vonstein, Veronika

```

/ APPLICANT: Demirjian, David
/ APPLICANT: Casadaban, Malcolm
/ TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
/ FILE REFERENCE: 95-963-H
/ CURRENT APPLICATION NUMBER: US/09/058.260B
/ CURRENT FILING DATE: 1999-04-10
/ EARLIER APPLICATION NUMBER: 60/001,995
/ EARLIER FILING DATE: 1996-08-07
/ EARLIER APPLICATION NUMBER: 60/009,704
/ EARLIER FILING DATE: 1996-01-11
/ EARLIER APPLICATION NUMBER: 60/019,580
/ EARLIER FILING DATE: 1996-06-12
/ EARLIER APPLICATION NUMBER: 08/694,078
/ EARLIER FILING DATE: 1996-08-06
/ EARLIER APPLICATION NUMBER: 08/781,802
/ EARLIER FILING DATE: 1997-01-10
/ EARLIER APPLICATION NUMBER: 08/827,810
/ EARLIER FILING DATE: 1997-04-11
/ NUMBER OF SEQ ID NOS: 37
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 31
/ LENGTH: 1756
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
/ OTHER INFORMATION: Gene from bacteria E027
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (130)..(1632)
/ US-09-058-260-31

```

Query Match	2.8%;	Score 39.2;	DB 3;	Length 1756;
Best Local Similarity	56.1%;	Pred. No. 0.26;		
Matches 74;	Conservative 0;	Mismatches 58;	Indels 0;	Gaps 0;
QY	205	TCACGTACATGGGATTCGCTTTGCGGGGCTCGGATGGCTTCGGTTCGTGTGAGAGCCA	264	
Db	26	TCCGGTGCATGGGATTCGCTTCAGGGGAACCTTTTAAACACTTGGATTGCAACCACTCCT	85	
QY	265	ACAACTTTTAAAGAAATATATAAAATAAAAAATAAAAAAATTTAAAGTGTTCGGAAGTGAATA	324	
Db	86	TAATCATTTAAGATTTAAATCGAAATTTAAATAATAATCAAAACAGCTGATTCAAATGAATA	145	
QY	325	CGGTGGGCACAG	336	
Db	146	CGTTCGTGGAAT	157	

RECORD 13
US-09-058-260-23
Sequence 23, Application US/09058260B
Patent No. 6218167
GENERAL INFORMATION:
APPLICANT: Allen, Larry
APPLICANT: Aikens, John
APPLICANT: Fonstein, Michael
APPLICANT: Vonstein, Veronika
APPLICANT: Demitjian, David
APPLICANT: Casadaban, Malcolm
TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
FILE REFERENCE: 95-963-H
CURRENT APPLICATION NUMBER: US/09/058,260B
CURRENT FILING DATE: 1999-04-10
EARLIER APPLICATION NUMBER: 60/001,995
EARLIER FILING DATE: 1996-08-07
EARLIER APPLICATION NUMBER: 60/009,704
EARLIER FILING DATE: 1996-01-11
EARLIER APPLICATION NUMBER: 60/019,580
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: 08/694,078
EARLIER FILING DATE: 1996-08-08
EARLIER APPLICATION NUMBER: 08/781,802

```

; EARLIER FILING DATE: 1997-01-10
; EARLIER APPLICATION NUMBER: 08/827,810
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1776
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:cloned esterase
; OTHER INFORMATION: Gene from bacteria E015
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (128)..(1630)
; US-09-058-260-23

Query Match      2.8%; Score 39.2; DB 3; Length 1776;
Best Local Similarity 56.1%; Pred.No.0.26;
Matches 74; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy      205  TCAAGTACATGTGAGTTCGCCCTTGTTCGGGGGCTTCGGATGGCTTGGGTCGTTGTGAGAAGCCA 264
Db      24   TCCGGTTCGATGAGTTCCTTCASGGGAACCTTTTAAACACCTTGAGTTTGACAAACCACTCT 83
Qy      265  ACAATTATTAAGAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 324
Db      84   TAATCATTTAAGATTTAAATGAAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 143
Qy      325  CGGTGGGGCAGA 336
Db      144  CGTTGGTGGAAA 155

```

```

RESULT 14
US-09-058-260-5
Sequence 5, Application US/09058260B
Patent No. 6218167
GENERAL INFORMATION:
APPLICANT: Allen, Larry
APPLICANT: Aikens, John
APPLICANT: Fonstein, Michael
APPLICANT: Vonstein, Veronika
APPLICANT: Demitrijan, David
APPLICANT: Casadaban, Malcolm
TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
FILE REFERENCE: 95-963-H
CURRENT APPLICATION NUMBER: US/09/058,260B
CURRENT FILING DATE: 1999-04-10
EARLIER APPLICATION NUMBER: 60/001,995
EARLIER FILING DATE: 1996-08-07
EARLIER APPLICATION NUMBER: 60/009,704
EARLIER FILING DATE: 1996-01-11
EARLIER APPLICATION NUMBER: 60/019,580
EARLIER FILING DATE: 1996-06-12
EARLIER APPLICATION NUMBER: 08/694,078
EARLIER FILING DATE: 1996-08-08
EARLIER APPLICATION NUMBER: 08/781,802
EARLIER FILING DATE: 1997-01-10
EARLIER APPLICATION NUMBER: 08/827,810
EARLIER FILING DATE: 1997-04-11
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 1952
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: clone
OTHER INFORMATION: gene from bacteria E011
FEATURE:
NAME/KEY: CDS
LOCATION: (197)..(1699)

```

Query Match	2.8%;	Score 39.2;	DB 3;	Length 1952;
Best Local Similarity	56.1%;	Pred. No. 0.27;	58;	Indels 0;
Matches 74;	Conservative 0;	Mismatches 0;	Gaps 0;	
205	TCAAGTACATGGATCCCTTTTCGCGGGCTTCGATGGCTTCGGTTCGTGTGAGAGGCCA	264		
93	TCCGGTGGATGGATTCGCTTCAGGGGAACTTTTAAACACTTGAGTTTGACACCACTCCT	152		
265	ACAAATTTATAGAATAATATAAAATATAAAATAAAAAATTTTAAAGTGTTCGAGAGTGAAAA	324		
153	TAATCAATTAAGATTTAAATGAATAATTAAATAATAATCAAAAAGAGTGATTCAAATGAATA	212		
325	CGGTGGGCAGA	336		
213	CGTTGGTGGAAA	224		
Db				

RESULT 15
US-08-781-802-5
; Sequence 5, Application US/08781802
; Patent No. 5969121
; GENERAL INFORMATION:
; APPLICANT: ALLEN, Larry
; APPLICANT: AIKENS, John
; APPLICANT: FONSTEIN, Michael
; APPLICANT: VONSTEIN, Veronika
; APPLICANT: DEMIRJIAN, David
; APPLICANT: CASADABAN, Malcolm
; TITLE OF INVENTION: Stable Biocatalysts for Ester Hydrolysis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 S. Wacker Drive 32nd Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 197..1699
; OTHER INFORMATION: /note= "E011 sequence of longest
; OTHER INFORMATION: open reading frame; other possible start codons ATG/met4;
; OTHER INFORMATION: TTG/leu7; GTG/val8; GTG/val15; GTG/val36; ATG/met62"
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 197..1699
; US-08-781-802-5

```

	Query Match	2.8%	Score 39.2;	DB 2;	Length 4090;
	Best Local Similarity	56.1%	Pred. No. 0.39;		
	Matches 74;	Conservative	0;	Mismatches 58;	Indels 0; Gaps 0;
QY	205	TCAAGTACATTCGATTGCCCTTTGTCGGGCTTGGATGCTTGGGTT	CGTGGTTCGTGGAAGCCA	264	
Db	93	TCCGGTCCGATGGATTGCTTTCAGGGGAACCTTTAAACACTT	GAGTTTGACACACCCTCT	152	
QY	265	ACAATTTTAAAGAAATATATAAAATATAAAATATAAAATTTAAGTGT	TGGAGTGAATA	324	
Db	153	TAATCAATTAAGATTTAAATGAATAATATAATCAAAAGAGTGTAT	CAATGAATA	212	
QY	325	CGGTGGGCAGA	336		
Db	213	CGTTGGTGAATA	224		

Search completed: January 21, 2004, 22:19:44
Job time : 98.7682 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 17:11:45 ; Search time 472.03 Seconds
(without alignments)
10461.606 Million cell updates/sec

Title: US-09-936-869-2
Perfect score: 1401
Sequence: 1 aaactgcagcaaatatga.....ggaaattttgtgtacaaatc 1401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	49.6	3.5	22684	13	US-10-017-161-725
2	49.6	3.5	29221	12	US-10-292-798-635
3	45	3.2	519	10	US-09-878-574-4292
4	43.8	3.1	625	13	US-10-027-632-44092
5	43.8	3.1	625	13	US-10-027-632-44093
6	43.8	3.1	625	13	US-10-027-632-44094
7	43.8	3.1	625	13	US-10-027-632-44095
8	43.8	3.1	625	14	US-10-027-632-44092
9	43.8	3.1	625	14	US-10-027-632-44093
10	43.8	3.1	625	14	US-10-027-632-44094
11	43.8	3.1	625	14	US-10-027-632-44095
12	43.6	3.1	26320	13	US-10-034-650-13
13	43.6	3.1	176373	14	US-10-095-407-17
14	43.4	3.1	409	11	US-09-918-995-6789
15	43.4	3.1	13326	13	US-10-311-455-1685

c 16	43.2	3.1	1840	12	US-10-292-798-2003
c 17	43.2	3.1	1840	13	US-10-017-161-2359
c 18	43.2	3.1	18434	13	US-10-311-455-1979
c 19	43.2	3.1	3673778	13	US-10-312-841-1
c 20	43	3.1	201	9	US-09-864-761-18750
c 21	43	3.1	201	9	US-09-864-761-21744
c 22	43	3.1	395	9	US-09-864-761-5016
c 23	43	3.1	406	9	US-09-864-761-2003
c 24	43	3.1	3066	13	US-10-148-687-8
c 25	42.8	3.1	428	10	US-09-960-352-573
c 26	42.6	3.0	491	11	US-09-918-995-11962
c 27	42.6	3.0	753	13	US-09-814-353-2034
c 28	42.6	3.0	753	13	US-09-814-353-8380
c 29	42.6	3.0	8530	13	US-10-311-455-1727
c 30	42.4	3.0	143305	10	US-09-729-920-3
c 31	42.2	3.0	327	10	US-09-960-352-4630
c 32	42.2	3.0	7657	13	US-10-311-455-1995
c 33	42.2	3.0	7657	15	US-10-239-676-185
c 34	42.2	3.0	12237	13	US-10-311-455-2331
c 35	42.2	3.0	16509	13	US-10-311-455-1293
c 36	42	3.0	446	10	US-09-960-352-3400
c 37	42	3.0	1954	9	US-09-925-301-82
c 38	42	3.0	2812	15	US-10-198-846-9693
c 39	42	3.0	152331	14	US-10-095-407-16
c 40	41.8	3.0	642	11	US-09-866-050A-532
c 41	41.8	3.0	642	15	US-10-152-661-532
c 42	41.8	3.0	11577	13	US-10-311-455-414
c 43	41.6	3.0	172	9	US-09-864-761-25620
c 44	41.6	3.0	511	9	US-09-864-761-8959
c 45	41.6	3.0	6071	13	US-10-311-455-340

ALIGNMENTS

RESULT 1
US-10-017-161-725
; Sequence 725, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABEYATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 725
; LENGTH: 22684
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(22684)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(615)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (753)..(878)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17062)..(17263)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (17606)..(17723)
; FEATURE:
; NAME/KEY: CDS


```
QY 385 CAACCTGATTTCTGGACCTCGAATACAGATATATGTTGGTGAAGAAATAAAGGAAGAGTGG 444
Db 391 AAAAGTTAAAAAATAATAATAAAAAAATAAATTTGTAATAAAAAAATAAATGATAAATCA 332
QY 445 AGTGCATTTGAAATGAATGAGAGCGCCACAAATGAGAGCGAATGAATGAATGAATATAT 504
Db 331 ATTAAAGAAATAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 272
QY 505 GCAAGAGTGCATTTCCCTATTATTTCAGAAATGTAT 541
Db 271 TAAAAATTATAATGATAAATAATTGAAAGAAATAAT 235
```

RESULT 4

```
US-10-027-632-44092
; Sequence 44092, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44092
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(625)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-44092
```

```
Query Match 3.1%; Score 43.8; DB 13; Length 625;
Best Local Similarity 57.8%; Pred. No. 0.32;
Matches 78; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
```

```
QY 256 GAGAAGCCACAAATTTATAGAATAATATAAATAAATAAATAAATAAATAAATAAATAA 315
Db 349 GAGAGAGACCCCTGCTCTCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 408
QY 316 AAGTGAATAACGGTGGCGAGAAATATACAGAGAGTACTTTTAAACAATGCGCAACCAAG 375
Db 409 AAATAAAGGATAGACAAATAGATCAATTGAGAGTCCATAAACAACCAACAAT 468
QY 376 GCAGATTCACAACCTT 390
Db 469 GCAGTCAAAATAGCTT 483
```

RESULT 5

```
US-10-027-632-44093
; Sequence 44093, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44093
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(625)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-44093
```

Query Match

```
3.1%; Score 43.8; DB 13; Length 625;
Best Local Similarity 57.8%; Pred. No. 0.32;
Matches 78; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
```

```
QY 256 GAGAAGCCACAAATTTATAGAATAATATAAATAAATAAATAAATAAATAAATAAATAA 315
Db 349 GAGAGAGACCCCTGCTCTCTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 408
QY 316 AAGTGAATAACGGTGGCGAGAAATATACAGAGAGTACTTTTAAACAATGCGCAACCAAG 375
Db 409 AAATAAAGGATAGACAAATAGATCAATTGAGAGTCCATAAACAACCAACAAT 468
QY 376 GCAGATTCACAACCTT 390
Db 469 GCAGTCAAAATAGCTT 483
```

RESULT 6

```
US-10-027-632-44094
; Sequence 44094, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
```

```
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44094
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(625)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-44094

Query Match          3.1%; Score 43.8; DB 13; Length 625;
Best Local Similarity 57.8%; Pred. No. 0.32;
Matches 78; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 256 GAGAGCCCAAACTTTAAGAATATATACAGAGAGTACTTTAACAATCGGCAACCAAG 315
    |||
Db 349 GAGAGACCCCTGCTCTAAATAAAAAATATAAATACAAATAGATCAATTGAAGAGTCCATTAACCAACCAACCAAT 468

QY 316 AAGTGAACCGTGGCGCAGAAATATACAGAGAGTACTTTAACAATCGGCAACCAAG 375
    |||
Db 409 AAATAAAAGGATAGACAAATAGATCAATTGAAGAGTCCATTAACCAACCAACCAAT 468

QY 376 GCAGATTCACAACTT 390
    |||
Db 469 GCAGTCAAAATAGCTT 483

RESULT 7
US-10-027-632-44095
; Sequence 44095, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-04-30
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44095
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(625)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-44095

Query Match          3.1%; Score 43.8; DB 14; Length 625;
Best Local Similarity 57.8%; Pred. No. 0.32;
Matches 78; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 256 GAGAGCCCAAACTTTAAGAATATATACAGAGAGTACTTTAACAATCGGCAACCAAG 315
    |||
Db 349 GAGAGACCCCTGCTCTAAATAAAAAATATAAATACAAATAGATCAATTGAAGAGTCCATTAACCAACCAACCAAT 408

QY 316 AAGTGAACCGTGGCGCAGAAATATACAGAGAGTACTTTAACAATCGGCAACCAAG 375
    |||
Db 409 AAATAAAAGGATAGACAAATAGATCAATTGAAGAGTCCATTAACCAACCAACCAAT 468

QY 376 GCAGATTCACAACTT 390
    |||
Db 469 GCAGTCAAAATAGCTT 483

RESULT 8
US-10-027-632-44092
; Sequence 44092, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-04-30
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44092
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(625)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-44092

Query Match          3.1%; Score 43.8; DB 14; Length 625;
Best Local Similarity 57.8%; Pred. No. 0.32;
Matches 78; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 256 GAGAGCCCAAACTTTAAGAATATATACAGAGAGTACTTTAACAATCGGCAACCAAG 315
    |||
Db 349 GAGAGACCCCTGCTCTAAATAAAAAATATAAATACAAATAGATCAATTGAAGAGTCCATTAACCAACCAACCAAT 408

QY 316 AAGTGAACCGTGGCGCAGAAATATACAGAGAGTACTTTAACAATCGGCAACCAAG 375
    |||
Db 409 AAATAAAAGGATAGACAAATAGATCAATTGAAGAGTCCATTAACCAACCAACCAAT 468

QY 376 GCAGATTCACAACTT 390
    |||
Db 469 GCAGTCAAAATAGCTT 483

RESULT 9
US-10-027-632-44093
; Sequence 44093, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-04-30
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 44093
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(625)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-44093

Query Match          3.1%; Score 43.8; DB 13; Length 625;
Best Local Similarity 57.8%; Pred. No. 0.32;
Matches 78; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 256 GAGAGCCCAAACTTTAAGAATATATACAGAGAGTACTTTAACAATCGGCAACCAAG 315
    |||
```



```

; OTHER INFORMATION: n = A,T,C or G
US-10-095-407-17

Query Match          3.1%; Score 43.6; DB 14; Length 176373;
Best Local Similarity 48.8%; Pred.No.10; Indels 0; Gaps 0;
Matches 118; Conservative 0; Mismatches 124;

QY   274 AGAATAATATAAAATAAAATAAAATAAAATTAAAGTGTGGAACTGAAAACGGTGCGGC 333
Db   122513 AAGAAGAGAAAAAAGAAAGAAAGAAAGAAAGATGGAAGGAAAGAAAGAAAGAA 122454

QY   334 AGAAATATACACAGAAGAGTACTTTAACCAATGCCAACCCAGGCGAGATTTCACACTTCA 393
Db   122453 GAATAGTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 122394

QY   394 TTCTGGACCTCGAATACGAGATTAATGCTGTAAGAAATTAAGCAAGAGTGGAGTCCATTT 453
Db   122393 AGAAAGAGAAAGAAAGGAGAGAAAGAAAGAAAGGAAAGAAAGAAAGAAAGCAACA 122334

QY   454 GAAATGAATGGAGAGCGCACAAAATGGAGGACGAAATTAATGAAATATAATGCAAGAGTG 513
Db   122333 GMAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 122274

QY           514 CA 515
Db   122273 AA 122272

RESULT 14
US-09-918-995-6789
; Sequence 6789, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseg, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6789
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(409)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-6789

Query Match          3.1%; Score 43.4; DB 11; Length 409;
Best Local Similarity 52.9%; Pred.No.0.32;
Matches 92; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY   151 GCTTGCTTCTGGTAATATGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGT 210
Db   164 GTTTGGTTTGGTTTAGAGAGGAATGAATAAGAGATTTGTTTAACTACCCCTAGG 223

QY   211 ACATTGGATTGCCCTTTGTCTGGGCTTGGATGCTTGGTGTGCGTGAGAGCCCAACAA 270
Db   224 AAAATGCTGCACATAAATAAAGGTTACTGCAGGGGTTTTCCCCTACACAGCANAAGAAA 283

QY   271 TATAAGAAATATATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 324
Db   284 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 337

RESULT 15
US-10-311-455-1685
; Sequence 1685, Application US/10311455

```

```

; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1685
; LENGTH: 13326
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1685

Query Match          3.1%; Score 43.4; DB 13; Length 13326;
Best Local Similarity 53.2%; Pred. No. 2.6;
Matches          92; Conservative          0; Mismatches          81; Indels          0; Gaps          0;

282 TATAAATAAAAAATTAAGTGTTCGAAGTGAAAACGGTGGCGGCAAGAATAT 341
9237 TTTTAAAAATGAAGAGAAATAAATTATATGTTTGTATGGTAGAGGATAGGAGAGAATTT 9296

342 ACACAGAAGAGTACTTTTAAACAATGCCAACCAAGGCAGATTCCAACTTCATTCTGGAC 401
9297 TTAGAGAAATAGAAATTTTGGCATCAAGTAGGAGTGNAGTTATTATAAATTAAGTGTT 9356

402 CTCGAATCAGAGATAATGGTGGTGAAGAAATAAGGAAGAGTGGAGTGCAATTG 454
9357 ATAGAGAAAAAGAAAAATGTAGGGATGAGATTGGGAAAGATATATGTGTTTTG 9409

```

Search completed: January 22, 2004, 01:20:35
Job time : 475.03 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 21, 2004, 17:02:36 ; Search time 1813.1 Seconds
(without alignments)
12185.050 Million cell updates/sec

Title: US-09-936-869-3
Perfect score: 909
Sequence: 1 atgggcgtgggaagattga.....cgcatttatcttctgtg 909

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estnu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_hic.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phs.*
27: em_gss_vri.*
28: gb_gss1.*
29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	310.2	34.1	697	CD482321	CD482321 at01-19m
2	278.2	30.6	823	CB973770	CB973770 CAB30003
3	276.8	30.5	819	CB975703	CB975703 CAB30007
4	275.4	30.3	668	CB009226	CB009226 VVC05C11

5	274.6	30.2	665	14	CB005658	CB005658 VVC024D03
6	273.8	30.1	785	14	CB969491	CB969491 CAB10002
7	273.8	30.1	843	14	CB969483	CB969483 CAB10002
8	271	29.8	754	12	BI311053	BI311053 EST531280
9	267	29.4	608	9	AW278878	AW278878 sf99n09.y
10	264.2	29.1	822	14	CB974938	CB974938 CAB30005
11	262.6	28.9	824	14	CB975572	CB975572 CAB30007
12	260	28.6	858	14	CA857928	CA857928 EST635183
13	258.6	28.4	615	13	BU045147	BU045147 PP_LEA002
14	258.6	28.4	631	13	BU043610	BU043610 PP_LEA001
15	257.8	28.4	649	9	AW184799	AW184799 se82f12.y
16	257.2	28.3	610	13	BU046256	BU046256 PP_LEA002
17	256.6	28.2	866	10	BG441292	BG441292 GA_EA001
18	255.2	28.1	634	14	CA902463	CA902463 PCS02695F
19	254.6	28.0	715	12	BI933168	BI933168 EST55305F
20	252.8	27.8	892	10	BG444639	BG444639 GA_EA002
21	251.6	27.7	605	13	BU045737	BU045737 PP_LEA002
22	251.4	27.7	761	9	AI731368	AI731368 BNIGH1937
23	250.8	27.6	776	10	BG445079	BG445079 GA_EA002
24	249.8	27.5	754	9	AI731375	AI731375 BNIGH1938
25	247.6	27.2	603	14	CB911729	CB911729 VVD135A10
26	247.6	27.2	605	14	CB917153	CB917153 VVD015E04
27	246.2	27.1	548	10	BE434089	BE434089 EST405167
28	246	27.1	538	10	BE433121	BE433121 EST399650
29	245.2	27.0	607	9	AW708451	AW708451 sk49C05.y
30	245	27.0	624	14	CB921120	CB921120 VVD050C09
31	244.4	26.9	663	14	CB918484	CB918484 VVD033H03
32	244.4	26.8	707	9	AI898432	AI898432 EST267875
33	243.2	26.8	665	9	AI487420	AI487420 EST245742
34	241.4	26.6	697	14	CB290594	CB290594 UCRCS01.0
35	241.4	26.6	632	13	BU048762	BU048762 PP_LEA003
36	241.2	26.5	686	13	BQ411600	BQ411600 GA_EQ004
37	241.2	26.5	656	13	BQ411583	BQ411583 GA_EQ004
38	239.6	26.4	698	10	BE659915	BE659915 7-H4 Gmax
39	239.6	26.4	537	9	AI486290	AI486290 EST244611
40	237.2	26.1	539	9	AI771796	AI771796 EST252896
41	237.2	26.1	539	9	AI484747	AI484747 EST243008
42	236.4	26.0	618	9	CB066648	CB066648 EST645271
43	236.4	26.0	729	14	CB066648	CB066648 EST531284
44	235	25.9	624	9	AW277843	AW277843 sf87d01.y

ALIGNMENTS

RESULT 1
CD482321
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD482321
at01-19msl-a12
CD482321
CD482321
EST.
Amborella trichopoda
Amborella trichopoda
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; stem Magnoliophyta; Amborellaceae;
Amborella.
1 (bases 1 to 697)
dePamphilis,C., Soltis,D., Soltis,P., Farmerie,W., Ma,H., Tanksley
S., Leebens-Mack,J., Field,D., Buzgo,M., Kim,S., Ilut,D., Landherr
L., Hu,Y., Wall,K., Albert,V., Carlson,J., Doyle,J., Frohlich,M.,
Miller,W., Oppenheimer,D. and Theissen,G.
Generation of ESTs from early flower buds of Amborella trichopoda
Unpublished
Contact: Claude dePamphilis or James Leebens-Mack
Penn State University
208 Mueller Laboratory, Department of Biology, ATTN Rm212, Penn
State University, University Park, PA 16802, USA
Tel: 814 863 6413
Fax: 814 865 9131

697 bp
linear
EST 04-JUN-2003
at01-19msl-a12
Amborella trichopoda
cdna clone
5', mRNA sequence.

149 TTCTGCAAGCAAGAAATGGGCTTTTGAAGAAGCTTATGAATATCAGTGTATGTAT 208
121 GCAGAACTGGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTATGAATTTGCAACAC 180
209 GCAGAACTTGGCCCTCATCGTCTTCTCCAGCGGGTGGTCTATGATCTATGACTCAACAC 268
181 AGCGTGAAGAGGACGATGAGAGTCAAGAGACTTGGTTCGACAAACACACGAGGG 240
269 AACATAAATCAACCATAGATAGTCAAGAGGCGCCAGCTCAGATAGTACAAATGAGGC 325
241 GCGATATCAGAGTCCAAATCTTCAGTATTGGCAACAGGAGGCTGTAAACTCAGACACAG 300
329 TCTACCATGAGATCAATGCCCATATATCAGCAGAGATCAGAAAGCTGCGCCAGCAA 388
301 ATTGACATTTTGCAAAATGCAATAGGATTTGATGGGTGACGGGTTCAGCTTTGAAC 360
389 ATACAGATGCTGCAGAAATCTTAACAGGCACTTAATGGGTGATTCCTTGGCTTCTTGA 448
361 ATTAAGCACTCAGCACTGAGTTGCTGCTTGAATGCTGAAAGGATCAGCGAGTGCATCC 420
449 GTGAGAGCTTAAGCAGCTCGAGAACAGGCTTGTAGGAGCATCAAGAAATCAGGTCG 508
421 AAAAAGAACAGATGTTGCTTGAAGATCGACATCATGCAAGAGAGGGAACACATATT 480
509 AAGAGCATGATGTTGCTTGGCTGAGATTGAATCTTGCAGAAAGGGAATTTGAGCTG 568
481 ATCCAGAGAAATGAGATTTCTGCAAGATAGCGAGTGTGAGATAGCCACACACAG 540
569 GAAATGAAAGCGCTATATCTCCGAACCAAGATTGCAAGAGTGGAGAGGCTTCAGCAAGCA 628
541 AACATGTTATCAGCTCCGGAATATGATGCACTGCCCGCATTCGACTCTCGAAATTTCCCTA 600
629 AACATGTTATCAACACATGATGTTCAATGCTATCCAGGCAATAGTTTCTGCAATTTCTTT 688
601 CATGCAAAATTAATCGATG 619
689 CAGCCCAATATGATTGAGG 707

RESULT 3
LOCUS CB975703 819 bp mRNA linear EST 01-MAY-2003
DEFINITION CAB30007_Ib_Fb_D07 Cabernet sauvignon Berry Stage I - CAB3 Vitis vinifera cDNA Clone CAB30007_Ib_Fb_D07 5', mRNA sequence.
ACCESSION CB975703
VERSION CB975703.1 GI:30298909
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
REFERENCE 1 (bases 1 to 819)
AUTHORS Goes da Silva, F., Iandolino, A., Lim, H., Baek, J., Jones, K. and Cook, D.
TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon' berries at various developmental stages
JOURNAL Unpublished
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq Primer: ACGTACCGACATATGCC.
Location/Qualifiers
1. 819
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet sauvignon"
/db_xref="taxon:29760"

/clone="CAB30007_Ib_Fb_D07"
/sex="Hermaphrodite"
/dev_stage="Berry stage I"
/lab_host="DHSalpa"
/clone_lib="Cabernet Sauvignon Berry Stage I - CAB3"
/note="Organ: Berry; Vector: pDNR; Site: 1; SfiI; Site 2: SfiI; CAB3 is a cDNA library of Vitis vinifera 'Cabernet Sauvignon' Clone 8 berries. Samples were collected after berry set from field-grown vines during stage I of berry growth, 17 days after full bloom. The average berry size was 6 millimeters. Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGTGTATCAACGAGATGCGCATACGCGCGG-3' and 5'-ATTAGAGCGGCGGCGCATG-DT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."
BASE COUNT 262 a 171 c 194 g 131 t 1 others
ORIGIN
Query Match 30.5%; Score 276.8; DB 14; Length 819;
Best Local Similarity 65.6%; Pred. No. 2.3e-65;
Matches 404; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
QY 4 GGTCTGGAGATGAGATAAGAGGATTCGAATTAATCTACGACCGACAGTCTCATTTC 63
Db 72 GGGAGAGAAAGATCGAGATCAAGAGGATCGAAACACGACCAACCGTCAGTCTCATATC 131
QY 64 TGCAGCGCCCGAAATGTTTATTAAAGAGCGGTATGAATTTATCAGTTCTTTTGTATGCA 123
Db 132 TGCAAGCGAAGGATGGGCTTTTGAAGAAGGCTTATGAATTTATCAGTCTATGTATGCA 191
QY 124 GAAGTGGCCCTCATCGCTTCTCCAGCAGAGGGAGACTTTATGAATTTGCCAACACACAGC 183
Db 192 GAAGTTGCCCTCATCGCTTCTCCAGCGCGGTGAGTCTATGAGTACTCAAAACAACAAAC 251
QY 184 GTGAAGAGGACGATTGAGAGGTACAAGAAGACTTTCGTTGACAAACACACCGAGGGGCG 243
Db 252 ATAAATCAACCATAGATAGGTACAAGAAGCGCAGCTCAGATAGTACAAATGGAGGCTCT 311
QY 244 ATATCAGATGCCAATTTCTAGTATTGGCAACAGAGGCTGTAACTCAGACACAGATT 303
Db 312 ACCATGGAGATCAATGCCAATATTACCAAGAAATCAGCAAGAGTGGCGCCAGCAATA 371
QY 304 GACATTTTGCAAAATGCAATAGGACATTTGATGGTGCAGCGGCTTACAGCTTTTGAACATT 363
Db 372 CAGATGTCGAGATTTCTAACGAGCACTTAAATGGGTGATTCCTTGGCTTCTTGACTGTG 431
QY 364 AAGAACTCAAGCAACTTGGTTGAGTTTCGACTTGAAGAAAGAAATCAGCGAGTCCGATCCAAA 423
Db 432 AAGAGCTTAAAGCAGCTCGAGAACAGGCTTGAACGAGGCATCAACAAGATCAGGTCCGAAG 491
QY 424 AAGACGAGATGTTGCTTGAAGAGATCGACATCATGCAAGAGAGGAGGAAACATATTATC 483
Db 492 AAGCATGTTGCTTGGCTGAGATTTGATTTGAGAAAGGAAATTTGAGCTGGAA 551
QY 484 CAGGAGATGAGATTTCTTCGCAAGATAGCCGAGTGTTCAGAAATAGCCAAACACGAAAC 543
Db 552 AATGAAGCGTATATCTCCGAAACCAAGATTGCAAGTGGAGAGGCTTTCAGAAAGCAAC 611
QY 544 ATGTTATCAGCTCCGATATGATGCACTGCCCGCTTCGACTCTCGAAATTTCTTACAT 603
Db 612 ATGGTATCAACACATGAGTTCAATGCCATTCAGGCAATTTCTTCCGAAATTTCTTCCAG 671
QY 604 GCAAAATCTAATCGATG 619
Db 672 CCCAATATGATTGAGG 687

CB009226 668 bp mRNA linear EST 10-JAN-2003
RESULT 4
LOCUS CB009226
CB009226

DEFINITION VVC065C11_230854 An expressed sequence tag database for abiotic stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera cDNA clone VVC065C11 5, mRNA sequence.

ACCESSION CB009226

VERSION CB009226.1

KEYWORDS GI:27586531

SOURCE EST.

ORGANISM Vitis vinifera

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.

AUTHORS 1 (bases 1 to 668)

TITLE Cushman,J.C.

JOURNAL An expressed sequence tag database for abiotic stressed berries of Vitis vinifera var. Chardonnay

COMMENT Unpublished

Contact: Cushman JC

Department of Biochemistry

University of Nevada

MS200, Reno, NV 89557-0014, USA

Tel: 775-784-1918

Fax: 775-784-1650

Email: jcushman@unr.edu

PCR Primers

FORWARD: T3 20mer

BACKWARD: T7 21mer

Plate: 065 row: C column: 11

Seq primer: T3 20mer

High quality sequence stop: 668.

FEATURES

source

1..668

/organism="Vitis vinifera"

/mol_type="mRNA"

/db_xref="taxon:29760"

/clone="VVC065C11"

/tissue_type="berries"

/dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"

/clone_lib="An expressed sequence tag database for abiotic stressed berries of Vitis vinifera var. Chardonnay"

/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 221 a 137 c 144 t

ORIGIN

Query Match 30.3%; Score 275.4; DB 14; Length 668;

Best Local Similarity 65.6%; Pred. No. 5.1e-65;

Matches 402; Conservative 0; Mismatches 211; Indels 0; Gaps 0;

Qy 1 ATGGGTCGTGGGAAGATGAGATAAAGAGGATTGAAATATCTACGACCGACAGGTCACCT 60

Db 56 ATGGGAGAGAGAAAGATCGAGATCAAGAGGATCGAAACACGACCAACCGTCAGGTCACA 115

Qy 61 TTCTGCAAGCGCCGATGTTTATTAAAGAGGCGTATGAAATATTCAGTCTTTTGTGAT 120

Db 116 TTCTGCAAGCGAAGATGGGCTTTTGAAGAGGCTTATGAATATTCAGTCTATGTGAT 175

Qy 121 GCAGAAGTGGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTATGAATTTGCCAACCCAC 180

Db 176 GCAGAAGTGGCCCTCATCGTCTTCTCCAGCGCGGTGCGAGTCTATGATGACTCAAAACAA 235

Qy 181 AGCTGAAGAGGAGATGAGAGGTACAGAGACTTCGTTGACACACACACCGAGGG 240

Db 236 AACATAAAATCAACCATAGTAGTACAAAGAGGCCAGCTCAGATATGATCAAAATGGAGGC 295

Qy 241 CGGATATCAGAGTCCAAATTCCTCAGTATTGGCAAGAGGCGTGGTAACTCAGACACACAG 300

Db 296 TCTACCATGGAGATCAATGCCAATCTACTACAGCAAGATCAGCAAGCTGCCGACGAA 355

Qy 301 ATTGACATTTTGCAAAATGCAAAATGAGCATTTGATGGGTGACGGGCTTACAGCTTTGAAC 360

Db 356 ATACAGATGCTGCAGAAATCTTAAACAGGCACCTAATGGGTGATTCCTTTGGCTTCCTTGACT 415

Qy 361 ATTAAGGAACTCAAGCAACTTGAGGTTCCGACTTGAAGAGGAAATCAGCCGAGTGCATCC 420

Db 416 GTGAAGGAGCTAAAGCAGCTCGAGAACAGAGCTTGAACGAGGCATCAACAAGATCAGGTG 475

Qy 421 AAAAAGAACGAGATGTTGCTTGAAGAGATCGACATATCGAGAGAGGGAACACATACCT 480

Db 476 AAGAAGCATGATGTTGCTTGGCTGAGATTGAGTACTTGAGAAAGGGAATGAGCTG 535

Qy 481 ATCCAGGAGATGAGATTCTTCGAGCAAGATAGCCGAGTGTTCAGATAGCACAACACG 540

Db 536 GAAAATGAACGCTATATCTCCGAACCAAGATTGAGAGTGGAGAGGCTTCAGCAAGCA 595

Qy 541 AACATGTTATCAGCTCCGATATGATGATGATGATGATGATGATGATGATGATGATGAT 600

Db 596 AACATGTTATCAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655

Qy 601 CATGCAAAATCTAA 613

Db 656 CAGCCCAATATGA 668

RESULT 5

CB005658

LOCUS CB005658

DEFINITION VVC024D03_137486 An expressed sequence tag database for abiotic stressed berries of Vitis vinifera var. Chardonnay Vitis vinifera cDNA clone VVC024D03 5, mRNA sequence.

ACCESSION CB005658

VERSION CB005658

KEYWORDS EST.

SOURCE CB005658.1

GI:27582963

ORGANISM Vitis vinifera

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; Vitaceae; Vitis.

AUTHORS 1 (bases 1 to 665)

TITLE Cushman,J.C.

JOURNAL An expressed sequence tag database for abiotic stressed berries of Vitis vinifera var. Chardonnay

COMMENT Unpublished

Contact: Cushman JC

Department of Biochemistry

University of Nevada

MS200, Reno, NV 89557-0014, USA

Tel: 775-784-1918

Fax: 775-784-1650

Email: jcushman@unr.edu

PCR Primers

FORWARD: T3 20mer

BACKWARD: T7 21mer

Plate: 024 row: D column: 03

Seq primer: T3 20mer

High quality sequence stop: 665.

FEATURES

source

1..665

/organism="Vitis vinifera"

/mol_type="mRNA"

/db_xref="taxon:29760"

/clone="VVC024D03"

/tissue_type="berries"

/dev_stage="mixed; 8, 9, 11, 13, 15, 16 weeks daf"

/clone_lib="An expressed sequence tag database for abiotic stressed berries of Vitis vinifera var. Chardonnay"

/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 220 a 137 c 165 g 143 t

ORIGIN

Query Match 30.2%; Score 274.6; DB 14; Length 665;

Best Local Similarity 65.7%; Pred. No. 8.5e-65;

Matches 400; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

Qy 1 ATGGGTCGTGGGAAGATGAGATAAAGAGGATTGAAATATCTACGACCGACAGTCACT 60

Db 56 ATGGGAGAGGAAAGATCGAGATCAAGAGGATCGAAAAACAGCAACCGTCAGGTCA 115
Qy 61 TTCTGCAAGCGCCGAATGGTTTATTAAAGAGCGGTATCAATATCAGTTCTTTTGAT 120
Db 116 TTCTGCAAGCGGAGGAATGGGCTTTTGAAGAGCGTTTATGAATATCAGTGTATGTAT 175
Qy 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGGAGCTTTATGAATTTGCCAACCC 180
Db 176 GCAGAGTGGCCCTCATCGTCTTCTCCAGCGGGTGTGATCTATGATCTCAACAC 235
Qy 181 AGCGTGAAGAGGAGGATGAGAGTCAAGAGAGCTTGCCTTGAACACACACAGGAGG 240
Db 236 AACATAAAATCAACCATAGATAGGTACAGAGGCCAGCTCAGATATGATAAATGGAGG 295
Qy 241 GCGATATCAGAGTCCAATTTCTCAGTATTGCAACAGAGGCTGTGAATCTCAGACACAG 300
Db 296 TCTACCATGAGATCAATGCCCATACTACCGACAGATCAGCAAGCTGCGCCAGCAA 355
Qy 301 ATTGACATTTTGCATAATGCAATAGGATTTGATGGGTGACGGGTTCAGCTTTGAAC 360
Db 356 ATACAGATGTCGAGAAATTTAAACAGGCACTTAATGGGTGATTCCTTGGCTTCTGACT 415
Qy 361 ATTAAGCACTCAGCAACTTGAGTTCGACTTTGAAAAGGAATCAGCGAGTGCATCC 420
Db 416 GTGAAGAGCTTAAGCAGCTCGAACAAGGTTGACGAGGATCAAGAAATCAGGTG 475
Qy 421 AAAAAGAACAGATGTTTCTTGAAGATCGACATCATGAGAGAGGGAACATATCTT 480
Db 476 AAGAAGCATGATGTTGCTTGGTGTAGATTGAGTACTTGCAGAAAGGGAATTTGAGCTG 535
Qy 481 ATCCAGAGATGAGATTTCTCGCAGCAAGATAGCCAGTGTGAGATAGCCACACAG 540
Db 536 GAAATGAAGCGGTATATCTCCGAACCAAGATTGCAGAGTGGAGAGGTTTCAGCAAGCA 595
Qy 541 AACATGTTATCAGTCCGGATATGATGCACTGCGCCCATTCGACTCTCGAAATTTCCCTA 600
Db 596 AACATGTTATCAACATGATGATTCATGCTCAATGCCATCCAGGCATTTCTCGCAATTTCTT 655
Qy 601 CATGCAAT 609
Db 656 CAGCCCAAT 664

RESULT 6
CB969491 786 bp mRNA linear EST 30-APR-2003
LOCUS CAB10002_IIla_Fa_E09 Cabernet Sauvignon Flower Pre-bloom - CAB1
DEFINITION Vitis vinifera cDNA clone CAB10002_IIla_Fa_E09 5', mRNA sequence.
CB969491
ACCESSION CB969491.1 GI:30251758
VERSION EST
KEYWORDS Vitis vinifera
SOURCE Vitis vinifera

ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; Vitaceae; Vitis.
REFERENCE
AUTHORS Goes da Silva, F., Tandolino, A., Lim, H., Baek, J., Jones, K. and Cook
D.

TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages

JOURNAL Unpublished
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu

Seq primer: ACGGTACCGGACATATGCC.
FEATURES
source Location/Qualifiers
1..786
/organism="Vitis vinifera"

/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clones="CAB10002_IIla_Fa_E09"
/sex="Hermaphrodite"
/dev_stage="Pre-bloom"
/lab_host="DHSalphi"
/clone_lib="Cabernet Sauvignon Flower Pre-bloom - CAB1"
/note="Worgan: Flower - Pre-bloom; Vector: pDNR; Site 1:
Sfil; Site 2: Sfil; CAB1 is a cDNA library of Vitis
vinifera cv. 'Cabernet Sauvignon' Clone 8 flowers. Samples
were collected approximately eleven days before onset of
bloom (clusters at this stage were fully developed and
flowers with calyptres or caps still attached. Sampled
vines were located at the University of California, Davis,
Experimental Vineyard. cDNAs were made by oligo-dT priming
and directionally cloned. 5' and 3' adaptors were used in
cloning as follows:
5'-AAGCTGTGATCAGCAGAGTGGCATACGGCCGG-3' and
5'-ATTCTAGAGCGGAGCGCCGACATG-DT(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."

BASE COUNT 243 a 178 c 181 g 184 t

Query Match 30.1%; Score 273.8; DB 14; Length 786;
Best Local Similarity 64.9%; Pred. No. 1.5e-64;
Matches 442; Conservative 0; Mismatches 227; Indels 12; Gaps 2;

Qy 1 ATGGGTCTGTGGGAAGATTGAGATTAAGAGGATTGAAAATACTAGAAACCGCAGGTCACT 60
Db 16 ATGGGAAGGGGAGATCGATCAGCGGATCGAATACCAACACCACTAACCGCAGGTACT 75
Qy 61 TTCTCAGAGCGCGAAATGGTTTATTAAGAGGGGTATGAATATCATAGTTCTTTTGAT 120
Db 76 TTCTCAGAGCGCGCAAGCGCTTACTCAAGAGGCTATGAATATCTGTCTCTATGTGAT 135
Qy 121 CGAAGTGGCCCTCATCGTCTTCTCCAGCAGAGGGAGCTTTATGAATTTGCCAACCCAC 180
Db 136 CGAGAGGTAGCCCTCATCGTCTTCTCCAGCGCGCCCTCTACGATATGCCAACAC 195
Qy 181 AGCGTGAAGAGGACGATTCAGAGGTACAGAGAGCTTCGCTTGAACAAACACCGAGGG 240
Db 196 AGTGTTAATCAACAATTGAGAGGTATTAAGAGGCAAGCGCTGATTCCTCCAAACCGGG 255
Qy 241 GCGATATCAGAGTCCAATTTCTCAGTATTGGCAACAGGAGGCTGTGTAATCTCAGACACAG 300
Db 256 TCCGTTTCTGAAGCCCAATGCTCAGTTCACAGCAAGAAATCCTCCAAACTGCATCAGCAG 315
Qy 301 ATTGACATTTTCAAAATGCAATAGGCATTTGATGGGTGACGGGCTTACAGCTTTGAAC 360
Db 316 ATTGCTAATTGACAGATTCGACAGGCAATGCTGGGTGAGTCTCTGGGCTCTTGAAT 375
Qy 361 ATTAAGGAATCAAGCAACTTGTGAGTTCGATTTGAAGAGGAATCAGCGGAGTGCATCC 420
Db 376 TTCAAGGACCTCAAGAGCCTGGAGATTTGGTTAGAGAAAGGCATTAAGTAGAATCAGATCC 435
Qy 421 AAAAAGACGAGATCTGCTTGAAGAGATCGACATCATGCAAGAGAGGGAACACATATCTT 480
Db 436 AAAAAGATGAGCTGTTGTTGCTGAGATCGATATATGCAAAAGAGGGAATAGACTTG 495
Qy 481 ATCCAGGAGATGAGATTTCTCGCAGCAGAGATAGCCGAGTGTGAGATAGC---CACAC 537
Db 496 CATAATGATAACCACTATCTCCGACCAAGGATAGCCGAGATAGAGAGAAACGAAACAGCAG 555
Qy 538 ACGAACATGTTATCAGCTCCGGAATATGATGCACTGCCCGCA-----TTCCACTCT 588
Db 556 ATGAGCTTGATGCCAGGAGGAGCTAACTATGAACTCATGCACTCTCAGCAGTTCCACTCT 615
Qy 589 CGAAATTTCTACATGCAAAATCTAATCGATGCGGCCATCTACTATGCAATCAGGAACA 648
Db 616 AGAAACTATTTCCAACTAAATGGAAGTGAACCAAAATCAATCATATCTCTGCCAAGACCAA 675


```

QY 649 ACAACGCTTCAGCTTGGCTGA 669
Db 676 CCAGCTCTTCAATTAGTTAA 696

RESULT 7
CB969483
LOCUS
DEFINITION
Vitis vinifera cDNA clone CAB10002_IIifa_Fa_p11 5', mRNA sequence.
ACCESSION
CB969483
VERSION
CB969483.1 GI:30251750
KEYWORDS
EST.
SOURCE
Vitis vinifera
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; Vitaceae; Vitis.
REFERENCE
1 (bases 1 to 843)
AUTHORS
Goes da Silva,F., Iandolino,A., Lim,H., Baek,J., Jones,K. and Cook
,D.
TITLE
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
JOURNAL
Unpublished
COMMENT
Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: ACGGTACCGACATATGCC.
FEATURES
Location/Qualifiers
1..843
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAB10002_IIifa_Fa_p11"
/sex="Hermaphrodite"
/dev_stage="Pre-bloom"
/lab_host="DHEalpa"
/clone_lib="Cabernet Sauvignon Flower Pre-bloom - CAB1"
/notes="Organ: Flower - Pre-bloom; vector: pDNR; Site:
Sfil; Site 2: Sfil; CAB1 is a cDNA library of Vitis
vinifera cv. 'Cabernet Sauvignon' Clone 8 flowers. Samples
were collected approximately eleven days before onset of
bloom (clusters at this stage were fully developed and
flowers with calyptras or caps still attached. Sampled
vines were located at the University of California, Davis,
Experimental Vineyard. cDNAs were made by oligo-dT priming
and directionally cloned. 5' and 3' adaptors were used in
cloning as follows:
5'-AAGCAGTGGTATCAACGAGAGTGCCTTACGCGGG-3' and
5'-ATTCTAGAGCGGAGCGGCGGACATG-dt(30)NN-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."
BASE COUNT 265 a 188 c 191 g 197 t
ORIGIN
Query Match 30.1%; Score 273.8; DB 14; Length 843;
Best Local Similarity 64.9%; Pred No. 1.6e-64;
Matches 442; Conservative 0; Mismatches 227; Indels 12; Gaps 2;
QY 1 ATGGGTCTGGGAATTGAGATAAGAGGATTGAAATCTAGAACCGCAGGTCACT 60
Db 16 ATGGGAAGGGGAGAGATCGAGATCAAGCGGATCGAAACACACCCTAACCGCGAGTTACT 75
QY 61 TTCTCCAGCGCGGAAGTGTATTAAAGAGGGGTATGAATTATCATGTTCTTTGTGAT 120
Db 76 TTCTCCAGCGCGGCAAGCGTTACTCAAGAGGGCTATGAATTATCTGTCTCTATGAT 135
QY 121 GCAGAAGTGGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTATGAATTGCCAACAC 180

```

```

Db 136 GCAGAGGTAGCCCTCATCGTCTTCTCCAGCGCGCGCGCTCTACGAATATGCCAACAC 195
QY 181 AGCGTGAAGAGGACGATTGAGAGGTACAGAAAGACTTGGCTTGACAAACACGAGGAGG 240
Db 196 AGTCTTAATCAACAATTGAGAGGTATTAAGAAAGCAAGCGCTGATCTCTCTCAACACCGGG 255
QY 241 GCATATACAGATCCCAATTCTCAGTATTGGCAACAGAGGCTGGTAAACTCAGACACAG 300
Db 256 TCCGTTTCTGAGCAATGCTCAGTTTACCAGCAAGAACTCTCCAAACTGCATCAGCAG 315
QY 301 ATTGACATTTTGCAAAATGCAAAATAGGCATTTGATGGGTGACGGGCTTACAGCTTTGAAC 360
Db 316 ATTGCTAATTTCAGAAATTCGAACAGGCACATGCTGGGTGAGTCTCTGGGCTCTCTGAAT 375
QY 361 ATTAAGGACTCAAGCAACTTGGCTTCGACTTGAAGAAAGAAATCAGCCGAGTCGATCC 420
Db 376 TTCAGGACCTTCAAGACCTTGGAGATTGGTTAGAGAAAGGCATTAAGTAGATCAGATCC 435
QY 421 AAAAAGAACACGAGATGTTGCTTGAAGAGATCGACATCATGACAGAGAGGGAACACATCTT 480
Db 436 AAAAAGAAATGAGTGTGTTGCTGAGATCGAGTATATGCAAAAGAGAGGAAATAGACTTG 495
QY 481 ATCCAGGAGATGAGATCTTCGACGACAGATCCGAGTGTCAAGATAGC---CACAAAC 537
Db 496 CATAATGATAACCGATATCTCCGAGCAAGGATAGCCGAGATGAGAGAGAGACAGCAG 555
QY 538 ACGAACATGTTTATCAGCTCCGGAATATGATGCACTGCCCGCA-----TTGACTCT 588
Db 556 ATGAGCTTGTATGCCAGGAGGAGCTAACTATGAACTCATGCCATCTCAGCAGTTGCACTCT 615
QY 589 CGAAATTTCTCATGCGAAATCTAATCGATCGGCGCCATCATCTATGACATCAGAAACAA 648
Db 616 AGAAACTATTTCCAACTAAATGGACTGCAACCAATATCATCTACTCTCGCCAGACCAA 675
QY 649 ACAACGCTTCAGCTTGGCTGA 669
Db 676 CCAGCTCTTCAATTAGTTAA 696

RESULT 8
BI311053
LOCUS
DEFINITION
EST5312803 GBSD Medicago truncatula cDNA clone pGBSD9f6 5' end,
mRNA sequence.
ACCESSION
BI311053
VERSION
BI311053.1 GI:14985380
KEYWORDS
EST.
SOURCE
Medicago truncatula (barrel medic)
ORGANISM
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
REFERENCE
1 (bases 1 to 754)
AUTHORS
Grusak,M.A., Samac,D.A., Town,C.D., Van Aken,S., Utterback,T., Cho
,J. and Fraser,C.M.
JSTs from developing reproductive tissues of Medicago truncatula
Unpublished
CONTACT: Michael A. Grusak
USDA/ARS Children's Nutrition Research Center
Baylor College of Medicine
1100 Bates Street, Houston, TX 77030-2600, USA
Tel: 713-798-7044
Fax: 713-798-7078
Email: mgrusak@bcm.tmc.edu
B397908e
TIGR sequence name: MTPAP27TK
More information is available at: www.medicago.org
Seq primer: SKmod (CTA GAA CTA gta gat CC).
FEATURES
Location/Qualifiers
1..754
/organism="Medicago truncatula"
source

```

```
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3980"
/clone="pGESD9F6"
/tissue_type="immature seeds"
/dev_stage="immature seeds, 11 to 19 days after
pollination"
/clone_lib="CESD"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:
XhoI; Immature seeds, collected from pods ranging in age
from 11 to 19 days after pollination, were harvested from
greenhouse-grown plants. Seeds were removed and
separated from pod walls and were immediately frozen in
liquid nitrogen. Seeds throughout the age range were
pooled for mRNA extraction. cDNA was prepared from polyA+
enriched RNA. The cDNA was directionally ligated using
the Unizap XR vector from Stratagene and packaged into
Gibapack III Gold packaging extracts. Plasmids containing
cDNA inserts were excised from the recombinant lambda-zap
phage using Ex-assist helper phage and propagated in
XLOLR cells."
BASE COUNT      274 a 121 c 159 g 200 t
ORIGIN
Query Match      29.8%; Score 271; DB 12; Length 754;
Best Local Similarity 63.0%; Pred. No. 8.9e-64;
Matches 418; Conservative 0; Mismatches 245; Indels 0; Gaps 0;
QY 1 ATGGTCTGGGAGATGATGAATGAAGAGGATTCGAAATACATACGACCGACAGGTCACCT 60
Db 25 ATGGGGAGGGGAAAGATTTGAGATTAAGAGGATTTGAGATTAACAAATAGGCAAGTGACT 84
QY 61 TTCTGCAAGCGCCGAATGGTTTATTAAAGAGCGGTATGAATATATCAGTTCTTTGTGAT 120
Db 85 TTTTGCAGAGAGGAGGATGACTTCGAGAGAGCTTATGAATATCAGTTTGTGTGAT 144
QY 121 GCAGATGGCCCTCATCGTCTTCGAGAGAGGAGACTTTATGAATTTGCCAACAC 180
Db 145 GCTGAAGTGTCTCATGTCTTCGAGCGGTGCGAGACTTTATGAATTTCAACAAC 204
QY 181 AGCGTGAGAGGAGGATTTGAGAGGTACAAGAGCTTGGTTCACACCAACACCGAGGG 240
Db 205 AACATAGGTCACAAATGATAGTACAAAAGCTTGTTCAGATCATCTCAAGTACCACC 264
QY 241 GCGATATCAGAGTCCAAATTCAGTATTTGGCAACAGAGGCTGGTAAATCAGACAAACAG 300
Db 265 ACTACCACTGAATTAACGCTCAGTATATCAGCAAGAACTGCAAAAGCTGAGACAACAA 324
QY 301 ATTGACATTTTGAAATGCAATAGGATTTGATGGTACCGGCTTACAGCTTTTGAC 360
Db 325 ATACAGATGCTGCAAAATTTCAAGGACCTTGTATGGTGTGATTAAGCACACTCACT 384
QY 361 ATTAAGGAACCTCAAGCAACTTGAGTTTCGACTTGAATAAGGAATTCAGCCGAGTCCATCC 420
Db 385 GTGAGGAATTTGAGCAGCTGGAGATAGACTTGAAGAGGAATCACTAGATTAGATCT 444
QY 421 AAAAAGACGAGATGTTGCTTGAAGAGATGACATCATGACAGAGAGGGAACATATCTT 480
Db 445 AAGAAACATGAGATGCTACTGGCTGAAATTTGAATPACTTTTCAAGAAAGGGAGATTGAGCTG 504
QY 481 ATCCAGGAGATGAGATCTTCGAGCAAGATAGCCGAGTGTGAGATAGCCCAACACG 540
Db 505 GAAATGAAATCTTTGCTTCGAGCTAAGATAATGATGTTGAGAGGCTTCCTCAAGTG 564
QY 541 AACATGTTATCAGTCCGGAATATGATGCACTGCCCGGATTCGACTCTCGAATTTCTTA 600
Db 565 AACATGTTTCTGGACACAGAGCTGAATGCAATCCAGCATTTAGCATCTCGTAATTTTTTC 624
QY 601 CATGCAATCTAATCGATGCGGCCCATCACTATCATCATCAGGAAACAAACAGCTTCAG 660
Db 625 AATCCAAACATGATGGAAGATGGTGAAATCTTCTTACCATCATCATGATAGAGATTTCTT 684
QY 661 CTT 663

/mol_type="mRNA" linear EST 02-DEC-2001
AW278878 608 bp mRNA
sf99h09.y1 Gm-cl019 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl019-4314 5' similar to TR:Q38836 Q38836 MADS-BOX PROTEIN
AGU11.1; mRNA sequence.
AW278878 608 bp mRNA
AW278878.1 GI:6667427
EST.
Glycine max (soybean)
Glycine max
Glycine max (soybean)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 608)
AUTHORS Shoemaker R., Keim P., Vodkin L., Erpeiding J., Coryell V., Khanna
A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J., Beck C.,
Wylie T., Underwood K., Steptoe M., Theising B., Allen M., Bowers
Y., Kitter E., Kohn S., Shin R., Gibbons M., Pape D., Harvey N., Schurk
R., Waterston R., and Wilson R.
Public Soybean EST Project
Public Soybean EST Project
Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 799 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 469.
Location/Qualifiers
1. 608
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3947"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl019-4314"
/tissue_type="immature seed coats of greenhouse grown
plants"
/lab_host="DH10B (Gibco BRL)"
/clone_lib="Gm-cl019"
/notes="Vector: pSPORT1 (Life Technologies); Site 1: Not I;
Site 2: Sal I; This cDNA library was constructed from mRNA
isolated from immature seed coats (200-300 mgs) of
greenhouse grown plants. The library was prepared using
the Life Technologies pSuperScript cDNA library
construction kit. Complementary DNA was synthesized from
mRNA using a poly (dt) sequence with a Not I restriction
site. Sal I linker adapters were ligated to the
blunt-ended cDNA fragments followed by Not I digestion.
The cDNA fragments were directionally cloned into the Not
I-Sal I restriction site of the pSPORT1 vector. The
ligated cDNA fragments were transformed into E.coli
electromax DH10B host cells (Gibco BRL). This library was
constructed by Dr. Lila Vodkin and Dr. Anu Khanna."
BASE COUNT      209 a 122 c 142 g 134 t 1 others
ORIGIN
Query Match      29.4%; Score 267; DB 9; Length 608;
Best Local Similarity 65.1%; Pred. No. 1e-62;
Matches 393; Conservative 0; Mismatches 211; Indels 0; Gaps 0;
QY 8 GTGGGAGATTGAGATAGAGGATTAAGATTAATACGACCGACAGGTCACCTTTCTGCA 67
```

Db 4 GGGGGAAGATCGAAATCAAAAGGATTGAGAACACAAATCGGCAAGTGACCTTCGCA 63
Qy 68 AGCGCCGAATGTTTATTAAAGAGCGGTATGAATTAATCACTTCTTGATGAGAG 127
Db 64 AGAAGAAATAGGCTTCGAAGAAAGCTTATAGCTGATGCTGCTGATGAGAG 123
Qy 128 TGCCCTCATGCTCTTCCAGCAGAGGAGACTTTATGAATTTGCCAACACACGCTGA 187
Db 124 TTGCCCTCATGCTCTTCCAGCGGTGGCGTCTCTATGATTTCCAAACAACATAA 183
Qy 188 AGAGGAGATTGAGAGGTACAGAAGACTTGGTTGACAAACACACGAGGCGGATAT 247
Db 184 GATCAACAATAGAGAGGTACAAAAGGCAATGTTCTGATCACTCAAGCGGAGACTACA 243
Qy 248 CAGAGTCCCAATTCCTAGTATTGGCAACAGAGGCTGGTAAACTCAGACACAGATTGACA 307
Db 244 CAGAAATCAATGCTCAGTATTATCAACAAGAACTCTCAAAAGCTGGACAGCAATACAGA 303
Qy 308 TTTTGCAAAATGCAATAGGCAATTTGATGGTGACGGGCTTACAGCTTTGACATTAAGG 367
Db 304 TGTGCAAAATTTAAACAGGCACTGATGGGTGATGCTTAAAGCACACTGATGTGAAGG 363
Qy 368 AACTCAAGCAACTTGAGGTTGCACTTCAAAAAGGAATCAGCCGAGTCCGATCCAAAAGA 427
Db 364 AACTTAAGCAGTTGGAGATAGACTTGAAGAGGAATCACTAGATCAGATCAAGATAC 423
Qy 428 ACAGATGTTGCTTGAAGAGATCGACATCATGAGAGAGAGGAGACATACTTATCCAGG 487
Db 424 ATGAGATGCTACTGCTGAAATTTGAATTTCCAGAAAAGGGAGATTGAACCTGGANAATG 483
Qy 488 AGAATGAGATTTCTCGCAGCAAGATAGCCGAGTGTGAGAAATAGCCCAACACAGCAATGT 547
Db 484 AGAATCTGCTCCCGAATTAAGATTAAGTCACTGACGTGAGAGGATTGACGAATTAACATGG 543
Qy 548 TATCAGCTCCGGAATATGATGCACTGCCCGCATTCGACTCTCGAAATTTCTTACATGCAA 607
Db 544 TATCTGGCCATGACTGATGATGCAATCCATGATGATGCTTCCGTAATCTTCTGCAATCCAT 603
Qy 608 ATCT 611
Db 604 ATAT 607

RESULT 10
CB974938 822 bp mRNA linear EST 01-MAY-2003
LOCUS CAB30005_Ib_Fb_E02 Cabernet Sauvignon Berry Stage I - CAB3 Vitis
DEFINITION vitifera cDNA clone CAB30005_Ib_Fb_E02 5', mRNA sequence.
ACCESSION CB974938
VERSION CB974938.1 GI:30298144
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; Vitaceae; Vitis.
REFERENCE 1 (bases 1 to 822)
AUTHORS Goes da Silva, F., Iandolino, A., Lim, H., Baek, J., Jones, K. and Cook
D.
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
Unpublished
CONTACT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: ACGTACCGGACATATGCC.
Location/Qualifiers
FEATURES
source
1..822

/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAB30005_Ib_Fb_E02"
/sex="Hermaphrodite"
/dev_stage="Berry stage I"
/lab_host="DHSalpha"
/clone_lib="Cabernet Sauvignon Berry Stage I - CAB3"
/note="Organ: Berry; Vector: pDNR; Site 1: SfiI; Site 2:
SfiI; CAB3 is a cDNA library of Vitis vinifera 'Cabernet
Sauvignon' Clone 8 berries. Samples were collected after
berry set from field-grown vines during stage I of berry
growth, 17 days after full bloom. The average berry size
was 6 millimeters. Sampled vines were located at the
University of California, Davis, Experimental Vineyard.
cDNAs were made by oligo-dT priming and directionally
cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AAGCAGTGGTATCAACGAGGAGGCGGCGG-3' and
5'-ATTCTAGAGCGGCGGCGGCGGCGG-3'. Library was
constructed using the Clontech Creator SMART kit and
size-selected to contain the 0.5-3 kb size fraction."
BASE COUNT 245 a 183 c 194 g 200 t
ORIGIN
Query Match 29.1%; Score 264.2; DB 14; Length 822;
Best Local Similarity 63.5%; Pred. No. 7.1e-62;
Matches 439; Conservative 0; Mismatches 243; Indels 9; Gaps 2;
Qy 1 ATGGGTGCTGGAGAGATTGAGATTAAGAGGATTGAAATTAATCACTAGCAACGACAGGTCACT 60
Db 97 ATGGGAAGAGGAGAGATCGAAATTAAGCGGATCGAGAACACCACCTAATTCGACAAGTCAACC 156
Qy 61 TTCTGCAAGCGCCGAAATGGTTTATTAAAGAGCGGTATGAATTAATCACTAGTCTTTTGTGAT 120
Db 157 TTCTGTAAGCGCCGCAACGGCTTCTCAAAAAGGCTATGAATTAATCGGCTTTATGAT 216
Qy 121 GCAGAGTGGCCCTCATGCTTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACAC 180
Db 217 GCAGAGTGGCTCTCATGCTTCTCCAGCGGTGGCGGCTTATGAATTAATGCTTAACAC 276
Qy 181 AGCGTGAAGAGCGAGTATGAGAGGTACAAGAGACTTGGCTTGACAAACACACCGAGGG 240
Db 277 AGTGTGAGACGACGATCGGAGGTACAAAAGTGTGTTCTGATTCCTCCATACCGGA 336
Qy 241 CGCATATCAGAGTCCAATTCAGTATTTGGCAAAGGAGGCTGGTAACTCAGACAACAG 300
Db 337 TCTGTTCTGAAGCTAATGCTCAGTTTTACCAAGAGAGGCTCCAAATTTGCGCGCTCAA 396
Qy 301 ATTGACATTTTGAAGATGCAATAGGCAATTTGATGGTGACGGCTTACAGCTTTGAAC 360
Db 397 ATCAGGATATACAGAAATTTGAAAGGCAATTTCTGGTGAGGCTCTTAGCTCTTGAAC 456
Qy 361 ATTAAGGAATCAAGCAACTTGAAGTTGCACTTGAAGAAAGAAATCAGCCGAGTGGATCC 420
Db 457 TTTAAGGAATCAAGAACTCGAGACCTGGTTGGAGAAAGGTATCAGCAGAATCAGATCC 516
Qy 421 AAAAGACGAGATGTTGCTTGAAGATGACATCATGAGAGAGGAGGACACATCTT 480
Db 517 AAAAGAAATGAATTTGCTGTTGCTGAAATCAGTATTTGAAAGAGGAGGATTTGAGCTC 576
Qy 481 ATCCAGGAGATGAGATTTCTTCGACGAGAGTAGCCGAGTGTCAAGATAGC---CACAAAC 537
Db 577 CAAAACACTAATTTCTCGAGACACAGATAGCTGAGAAATGAGAGAGGAGGAGGAGGAG 636
Qy 538 AGAAGATGTTTATCAGCTCCGGAATATGA-----TGCACTGCCGGGATTTGAGCTCTGA 591
Db 637 ATGAACCTGATGCCAGGCTCTCAGTACGAGAGCGTGGCGGAGAGCCATAGCTCTCAA 696
Qy 592 AATTTCTTACATGCAATCTAATCGATCGCGGCCCATCACTATGACATCAGGAGGAGGAGGAG 651
Db 697 AACTTGTCTCTCTGTTAACTCTCTGAGACCTAATCACCATTACTCTCTGCGCAGGAGGAGGAG 756

```

QY      652  ACGCTTCAGCTTGCTGAAAGCTTGAAGCGGT 682
Db      757  GCTCTCAACTAGTGTGAGAAGTGTGCTGT 787

RESULT 11
CB975572
LOCUS   CB975572               824 bp      mRNA      linear      EST 01-MAY-2003
DEFINITION   CAB30007_IVb_Fb_A06 Cabernet Sauvignon Berry Stage I - CAB3 Vitis
              vinifera cDNA clone CAB30007_IVb_Fb_A06 5', mRNA sequence.
ACCESSION   CB975572
VERSION     CB975572.1 GI:30298778
KEYWORDS    EST.
SOURCE      Vitis vinifera
            Vitis vinifera
ORGANISM    Vitis vinifera
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; Vitaceae; Vitis.
REFERENCE   1 (bases 1 to 824)
AUTHORS    Goes da Silva,F., Iandolino,A., Lim,H., Baek,J., Jones,K. and Cook
            D.
TITLE      Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
            berries at various developmental stages
JOURNAL    Unpublished
COMMENT    Contact: Douglas Cook, PhD
            CAES Genome Facility
            UC Davis Plant Pathology
            One Shields Ave, Davis, CA 95616, USA
            Tel: 530 754 6361
            Fax: 530 754 6617
            Email: drcook@ucdavis.edu
            Seq primer: ACGGTACCGACATATGCC.
            Location/Qualifiers
                1..824
                /organism="Vitis vinifera"
                /mol_type="mRNA"
                /cultivar="Cabernet Sauvignon"
                /db_xref="taxon:29760"
                /clone="CAB30007_IVb_Fb_A06"
                /sex="Hermaphrodite"
                /dev_stage="Berry stage I"
                /lab_host="DH5alpha"
                /clone_lib="Cabernet Sauvignon Berry Stage I - CAB3"
                /note="Organ: Berry; Vector: pDNR; Site 1: Sfil; Site 2:
                Sfil; CAB3 is a cDNA library of Vitis vinifera 'Cabernet
                Sauvignon' Clone 8 berries. Samples were collected after
                berry set from field-grown vines during stage I of berry
                growth, 17 days after full bloom. The average berry size
                was 6 millimeters. Sampled vines were located at the
                University of California, Davis, Experimental Vineyard.
                cDNAs were made by oligo-dT priming and directionally
                cloned. 5' and 3' adaptors were used in cloning as follows:
                5'-AAGCAGTGGTATCAACGACAGTGGCATTACGCCGGG-3' and
                5'-ATTCTAGAGCGCGGCGGACATG-dT(30)NN-3'. Library was
                constructed using the Clontech Creator SMART kit and
                size-selected to contain the 0.3-3 Kb size fraction."
BASE COUNT 241 a 195 c 184 g 204 t
ORIGIN

Query Match      28.9%; Score 262.6; DB 14; Length 824;
Best Local Similarity 63.4%; Pred No. 2e-61;
Matches 438; Conservative 0; Mismatches 244; Indels 9; Gaps 2;

QY      1  ATGGGTGCTGGAGATTGAGATAAGAGATTGAAAATCTACGAACCGCAGGTCAC 60
Db      113  ATGGGAAGAGGAAGATCGAAATTAAGCGGTATCGAAGACCACTAATCGAAGTCACC 172
QY      61  TTCTGCAGCGCGGAATGTTTATTAAAGCGCTATCAATTATCAGTCTTTGTGAT 120
Db      173  TTCTGTAGCCCGCAACGCTTCTCTCAAAAGGCTTATGAATTATCGTCTTATGTAT 232
QY      121  GCAGAAAGTGCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAATTCGCCAACAC 180

```

/db xref="taxon:3880"
/clone="pCLSD-26815"
/tissue_type="Immature seeds"
/dev_stage="25 to 35 days after pollination"
/lab_host="XLOLR"
/clone_lib="GLSD"
/notes="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Immature seeds, collected from pods ranging in age from 25 to 35 days after pollination, were harvested from greenhouse-grown plants. Seed were removed and separated from pod walls and immediately frozen in liquid nitrogen. Seeds throughout the age range were pooled for mRNA extraction. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XLOLR cells."
BASE COUNT 301 a 134 c 179 g 244 t
ORIGIN

Query Match 28.6%; Score 260; DB 14; Length 858;
Best Local Similarity 63.0%; Pred. No. 1.1e-60;
Matches 418; Conservative 0; Mismatches 245; Indels 1; Gaps 1;
QY 1 ATGGGTCGTGGAGATTGAGATAAGA-CGATTGAAATATCTACGACCGACAGCTCAC 59
Db 29 ATGGGAGGGGAGGATTGAGATAAAGGGATTGAGATACAAATAGGCAAGTGAC 88
QY 60 TTTCTGCAAGCGCGCAATGTTTATTAAAGAGGGCGGTATGAATATTCAGTTCTTTGTGA 119
Db 89 TTTTTCGAGAGAGGAGTGGACTTCTGAAGAAGCTTATGAATATTCAGTTCTTTGTGA 148
QY 120 TGCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTGAAATTTGCCAACCA 179
Db 149 TGTCTGAAGTTGCTCTCATCTGCTTCTCCAGCGGTGGCAGACTTTATGAATATTCAAACAA 208
QY 180 CAGCGTCAAGAGGAGGATTGAGAGGTACAGAGACTTGGTTGACAAACACACGAGG 239
Db 209 CAACATAGGTCAACAATTGATAGGTACAAAAGCTTGTTCAGATCATTCAGTACCAC 268
QY 240 GGCATATCAGAGTCCAAATTTCTAGTATTGGCAACAGAGCGTGGTAACTCAGACACA 299
Db 269 CACTTACCCTGAAATTAACGCTCAGTATTATCAGCAAGAACTCTGCAAGCTGAGACACA 328
QY 300 GATTGACATTTGCAAAATGCAATAGGCAATTTGATGGGTGACGGGCTTACAGCTTTGAA 359
Db 329 AATACAGATGCTGCAAAATTTCTACGCGCACTTGTGGGTGATGATTAAGCACTCAC 388
QY 360 CATTAAAGGAATCAAGCAACTTGGATTGCACTTGAAGAGGAATCAGCGAGTGGGATC 419
Db 389 TGTGAAGGAATTGAAGCAGCTGGAGATAGACTTTGAAGAGGAATCACTAGATTTAGATC 448
QY 420 CAAAAAGAACGAGATGTTGTTGAAGAGATCGAATCATGCGAGAGAGGAAACACATACT 479
Db 449 TAAGAAACATCAGATGCTTACTGGCTGAAATTTGAATCTTTTCAAGAAAGGAGATTGAGCT 508
QY 480 TATCCAGAGAAATGAGATTTCTTCAGCAAGATAGCGAGGTGTCAGAAATAGCCACACAC 539
Db 509 GGAATAATGAAATCTTTGCTTCGCACTAAGATTAATGATGTTGAGAGGCTTCTCAAGT 568
QY 540 GAACATGTTATCAGCTCGGAATATGATGCACTCCCGCAATTCGACTCTCGAAATTTCT 599
Db 569 GAACATGTTTCTGACAAAGAGCTGAATGCAATCCAGCAATTAGCATCTCGTAATTTTT 628
QY 600 ACATGCAATCTAATCGATCGGCGCCATCATCTATGCATCAGGAGCAACAAACAGCTTCA 659
Db 629 CAATCCAAACATGATGGAAGATGTTGAACTTCTTACATCATCAGATAGAGATTCT 688
QY 660 GCTT 663
Db 689 TCAT 692

RESULT 13
BU045147
LOCUS
DEFINITION
PP_LEA0021K08f Peach developing fruit mesocarp Prunus persica cDNA clone PP_LEA0021K08f, mRNA sequence.
BU045147
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Prunus persica (peach)
Prunus persica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
1 (bases 1 to 615)
Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.
Peach Model Genome for Rosaceae
Unpublished
Contact: Abbott, A.
Dept of Genetics and Biochemistry
Clemson University
122 Long Hall, Clemson University, Clemson, SC 29634, USA
Tel: 864 656 3060
Fax: 864 656 6879
Email: aalbert@clemson.edu
Total High Quality bases = 585
Seq primer: TAATACGACTCATATAGG
High quality sequence stop: 615.
Location/Qualifiers
1..615
/organism="Prunus persica"
/mol_type="mRNA"
/cultivar="Loring"
/db_xref="taxon:3760"
/clone="PP_LEA0021K08f"
/tissue_type="Mesocarp"
/lab_host="E. coli"
/clone_lib="Peach developing fruit mesocarp"
/note="Vector: pBluescript II SK(-); Site 1: EcoRI; Site 2: XhoI; authority=Prunus persica L. Batsh; The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis go to
http://www.genome.clemson.edu/projects/peach. To order this clone go to http://www.genome.clemson.edu/orders"
BASE COUNT 219 a 123 c 144 g 129 t
ORIGIN

Query Match 28.4%; Score 258.6; DB 13; Length 615;
Best Local Similarity 67.6%; Pred. No. 2.2e-60;
Matches 363; Conservative 0; Mismatches 174; Indels 0; Gaps 0;
QY 2 TGGTCTGGGAGATTGAGATTAAGAGGATTGAAATCTACGAAACCGACAGGTCACTT 61
Db 65 TGGTATGGGAGAGATCGAGATCAAGCGATCGAAACACACAAACCGTCAAGTCACT 124
QY 62 TCTCAAGCGCGCAATGTTTATTAAAGAGGGGTATGAATTAATCACTTCTTGTGATG 121
Db 125 TCTCAAAAGGCGCAATGGTTGCTCAAGAGGGCTATGAATCTCTCTACTCTGTGATG 184
QY 122 CAGAGTGGCCCTCATCTCTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACCA 181
Db 185 CAGAGTTGCTCTCATAGTTTTTCTTAACCGTGGCCGACTTTATGAGTATGCCAACACA 244
QY 182 GCGTCAAGAGGAGGATTGAGAGGTTAAGAGAGACTTGGCTTGACAAACACCGAGGGG 241
Db 245 GTGTTAAAGAAACAATTGAGAGGTACAAAGAGGATGCGCAGAGTCTCAAAATACCGAT 304
QY 242 CGATATCAGAGTCCCAATTTCTCAGTATTCGACACAGAGGCTGGTAACTTCAGACACAGA 301
Db 305 CTGTTTCCAGAGCTAGTACTCAGTACTACCAAGAGAGCTCGGAATCTGCGGCTCAGA 364

```

QY 302 TTGACATTTTGCRAAATGCAATAGGCAATTTGATGGGTGACGGCTTACAGCTTTTGAACA 361
Db 365 TAGGCAATTTGCAGAACTCAGCAGGCAATATGATGGCGAGTCACTTACGATTCATGANT 424
QY 362 TTAAGGAACCAACCAACTTGGATTCGACTTTGAAAGAAAGAAATCAGCCGAGTCGATCCA 421
Db 425 TGAAGACCTGGAAGAACTCGAGAGTAAACTAGAGAAAGGAATCAACAGAAATCAGATCCA 484
QY 422 AAAAGACGGAGATGCTGCTTGAACAGATCGACATCATGCGAGAGGGAACACATACCTTA 481
Db 485 AGAAGATGAACTTTGTTGCCGAATTTGATGATGCGAAGAGGAATTTGACTTGC 544
QY 482 TCCAGGAGAAATGAGATTTCTTGCAGCAAGATAGCCGAGTGTCAAGATAGCCACAACA 538
Db 545 ATAACAACAACCACTCTACGAGCAAGATAGCTGAGATGAGAGAGCCAGCAA 601

RESULT 14
BU043610
LOCUS
DEFINITION
PP LEA0016J23f Peach developing fruit mesocarp Prunus persica CDNA
BU043610_
BU043610.1 GI:22483687
EST.
SOURCE
Prunus persica (peach)
ORGANISM
Prunus persica
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Rosales; Rosaceae; Amygdaloideae; Prunus.
1 (bases 1 to 631)
Callahan,A., Palmer,M., Main,D., Wing,R. and Abbott,A.
Peach Model Genome for Rosaceae
Unpublished
Contact: Abbott, A.
Dept of Genetics and Biochemistry
Clemson University
122 Long Hall, Clemson University, Clemson, SC 29634, USA
Tel: 864 656 3060
Fax: 864 656 6879
Email: aalbert@clemson.edu
Total High Quality bases = 586
Seq primer: TAATACGACTCACTATAGGG
High quality sequence stop: 631.
Location/Qualifiers
1..631
/organism="Prunus persica"
/mol_type="mRNA"
/cultivar="Loring"
/db_xref="taxon:3760"
/clone="PP_LEA0016J23f"
/tissue_type="Mesocarp"
/lab_host="E. coli"
/clone_lib="Peach developing fruit mesocarp"
/note="Vector: pluscript II SK(-); Site 1: EcoRI;
Site 2: XhoI; authority=Prunus persica L. Batish; The
sequence has been trimmed to remove vector sequence and
contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis go to
http://www.genome.clemson.edu/projects/peach. To order
this clone go to http://www.genome.clemson.edu/orders"
BASE COUNT 222 a 124 c 153 g 132 t
ORIGIN

Query Match 28.4%; Score 258.6; DB 13; Length 631;
Best Local Similarity 57.6%; Pred. No. 2.2e-60;
Matches 363; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 2 TGGGTGCGGGAAGATTGAGATTAAGAGGATTTGAAATTAATCTACGAAACCGACGATCACTT 61
Db 64 TGGGTAGGGGAAGATCGAGATCAAGCGGATCGAAACACACAAACCGTCAGTCACTT 123

```

```

QY 62 TCTCAAGCGCCGAAATGTTTATTAAAGAGGGGTATGAATTATCAGTTCTTTGTGATG 121
Db 124 TCTCAAGAGCGCCAAATGGTTGCTCAAGAGGGCTATGAACCTCTCTGATCTGTGATG 183
QY 122 CAGAAAGTGGCCCTCATCGTCTTTCCAGCAGAGGAGACTTTTATGAATTTGCCAAACCA 181
Db 184 CAGAGGTTGCTCTCATAGTTTTTCTTAACCGTGGCCGACTTTTATGAGTATGCCAAACA 243
QY 182 GGGTCAAGAGGAGCAATTCAGAGGTACAAGAGACTTGGTTGACAAACAACCGAGGGG 241
Db 244 GTGTTAAGAAACAATTCAGAGGTACAAGAGGCATGCGCAGAGTCTACAATACCGAT 303
QY 242 CGATATCAGAGTCCAAATTTCTCAGTATTCGCAACAGAGAGCGTGGTAAATCTCAGACA 301
Db 304 CTGTTTCCGAAGCTAGTACTCAGTACTACCAACAAGAAAGCTCGGAAACTGCGCGTCA 363
QY 302 TTGACATTTTGCRAAATGCAATAGGCAATTTGATGGGTGACGGCTTACAGCTTTGAACA 361
Db 364 TAGGCAATTTGCAGAACTCAGCAGGCAATATGATGGCGAGTCACTTACGATTCATGA 423
QY 362 TTAAGGAATCAAGCAACTTGGAGTTTCGACTTTGAAAAAGGAATCAGCCGAGTCGATCCA 421
Db 424 TGAAGACCTGGAAGAACTCTGGAGAGTAACTAGAGAAAGGAATCAACAGAAATCAGATCCA 483
QY 422 AAAAGACGAGATGTTGCTTGAAGAGATCGACATCATGCGAGAGGGAACACATACCTTA 481
Db 484 AGAAGATGAACTCTTGTTCGCGAAATTTGAGTACATCGAAAAAGGGAATTTGACTTGC 543
QY 482 TCCAGGAGAAATGAGATTTCTTGCAGCAAGATAGCCGAGTGTCAAGATAGCCACAACA 538
Db 544 ATAACAACAACCACTCTACGAGCAAGATAGCTGAGATGAGAGAGCCAGCAA 600

RESULT 15
AW184799
LOCUS
DEFINITION
se85f12.y1 Gm-cl023 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl023-1224 5' similar to TR:Q38836 Q38836 MADS-BOX PROTEIN
AGL11.; mRNA sequence.
ACCESSION
AW184799
VERSION
AW184799.1 GI:6454186
KEYWORDS
EST.
SOURCE
Glycine max (soybean)
ORGANISM
Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 649)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk
,R., Ratter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 453.
Location/Qualifiers
1..649
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"

FEATURES
source

```

```
/clone="GENOME SYSTEMS CLONE ID: Gm-cl023-1224"
/tissue_type="seed coats of greenhouse grown plants"
/lab_host="DH10B"
/clone_lib="Gm-cl023"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; This
cDNA library was constructed from mRNA isolated from seed
coats (100-200mg) of greenhouse grown plants. The
library was prepared using the Life Technologies
pSuperscript cDNA library construction kit. Complimentary
DNA was synthesized from mRNA using a poly (dT) sequence
with a Not I restriction site. Sal I linker adapters
were ligated to the blunt-ended cDNA fragments followed by
Not I digestion. The cDNA fragments were directionally
cloned into the Not I-Sal I restriction site of the
pSPORT1 vector. The ligated cDNA fragments were
transformed into E.coli Electromax DH10B host cells. This
library was constructed by Dr. Lila Vodkin and Dr. Anu
Khanna."
```

BASE COUNT	231 a	123 c	156 g	139 t
ORIGIN				
Query Match	28.4%;	Score 257.8;	DB 9;	Length 649;
Best Local Similarity	65.2%;	Pred. No. 3.8e-60;		
Matches 379;	Conservative 0;	Mismatches 202;	Indels 0;	Gaps 0;
QY	1	ATGGGTCGTGGGAGATTGAGATAAGAGGATTGAAATTAATCTACGAACCGCAGAGTCACT	60	
Db	68	ATGGGAGGGGGGAGATCGAATCAAAAGATTGAGACACAAATCGCAAGTGACC	127	
QY	61	TTCTGCAAGCCCGCAATGGTTTATTAAAGAGCGGTATGAATTCAGTCTCTTTGTGAT	120	
Db	128	TTCTGCAAGAGAGAAATGGGCTTCTGAAGAAAGCTTATGAGCTGTCAGTGTGTGTGAT	187	
QY	121	GCAGAGTGGCCCTCATCTCTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACACAC	180	
Db	188	GCAGAGTTGGCCCTCATCTCTCTCCAGCGGTGGCGTCTCTATGAGTATTCACAAAC	247	
QY	181	AGCGTGAAGAGGAGATGAGAGTCAAGAGACTTGCCTTCAACACACACGAGGG	240	
Db	248	AACATAAGATCAACAATAGAGAGGTACAAAAGGCATGTTCTCATCACTCAAGCGGAGC	307	
QY	241	CGGATATCAGAGTCCAATTTCTCAGTATTGGCAAGAGGCTGTAACTCAGACACACAG	300	
Db	308	ACTACCAAGAAATCAATGCTCAGTATTATCAGCAGAAATCTGCAAGCTGCCACAGCA	367	
QY	301	ATTGACATTTTGCAAAATGCAATAGGATTTGATGGTGACGGGCTTACAGCTTTGAAC	360	
Db	368	ATACAGATGCTGCAAAATTTCTACAGGCACCTGATGGGTGATGCTTTAGCACACTGACT	427	
QY	361	ATTAGGACTCAGCAACTTGAGGTTCGACTTGAAGAGGAATCAGCCGAGTGGATCC	420	
Db	428	GTGAAGGAACCTTAAGCAGTTGGAGATAGACTTGAAGAGGAATCCTAGAAATCAGATCT	487	
QY	421	AAAAAGAACGAGATGTTGCTTTGAAGAGATCGACATCATGCGAGAGAGGAAACATACCTT	480	
Db	488	AAGAAACATGAGATGCTACTGGCTGAAATTCGAATCTCCAGAAAGGAGAGATTGAACTG	547	
QY	481	ATCCAGGAGATGAGATTTCTCGCAGCAGATAGCCGAGTGTCAAGATAGCCACACAG	540	
Db	548	GAAATGAAATCTTTTGGCTCCGAACCTAAGATAACTGACGTGGAGAGGATTTCAGCAAGTA	607	
QY	541	AACATGTTATCAGTCCCGAATATGATCACTGCCCGCATTT	581	
Db	608	AACATGGTTTCTGGGCCAGAACTGAATGCCATTCAGCATTT	648	

Search completed: January 22, 2004, 00:54:24
Job time : 1821.1 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 17:11:45 ; Search time 306.264 Seconds
(without alignments)
10461.606 Million cell updates/sec

Title: US-09-936-869-3

Perfect score: 909

Sequence: 1 atggctcggtgggaagattga.....cgcatgttatctttctgttg 909

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US09_PUB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	28.6	1219	15	US-10-104-580-10
2	259.8	28.6	723	15	US-10-104-580-11
3	255.6	28.1	959	10	US-09-978-730-5
4	255.6	28.1	959	10	US-09-978-729A-5
5	255.6	28.1	959	10	US-09-981-087A-5
6	255.6	28.1	959	10	US-09-978-382A-5
7	255.6	28.1	959	11	US-09-978-740A-5
8	248.2	27.3	1714	15	US-10-104-580-15
9	248.2	27.3	1159	15	US-10-104-580-14
10	245	27.0	896	10	US-09-978-730-3
11	245	27.0	896	10	US-09-978-729A-3
12	245	27.0	896	10	US-09-981-087A-3
13	245	27.0	896	10	US-09-978-382A-3
14	245	27.0	896	11	US-09-978-740A-3
15	230.2	25.3	905	15	US-10-286-264-160

16	201.8	22.2	795	13	US-10-259-165-775	Sequence 775, App
17	171.2	18.8	1280	10	US-09-970-624-1	Sequence 1, Appli
18	164.6	18.1	1038	11	US-09-819-142-3	Sequence 3, Appli
19	164.6	18.1	1038	11	US-09-334-455-31	Sequence 31, Appl
20	160.4	17.6	845	9	US-09-770-445-644	Sequence 644, App
21	157.8	17.4	1210	11	US-09-819-142-17	Sequence 17, Appl
22	157.8	17.4	1210	11	US-09-934-455-403	Sequence 403, App
23	157.8	17.4	1210	11	US-10-295-403-69	Sequence 69, Appl
24	152.8	17.0	824	13	US-10-259-165-161	Sequence 161, App
25	152.8	16.8	645	10	US-09-938-842A-1776	Sequence 1776, Ap
26	152.8	16.8	645	10	US-09-853-450-33	Sequence 33, Appl
27	148.6	16.3	959	11	US-09-819-142-13	Sequence 13, Appl
28	148.6	16.3	959	11	US-09-934-455-355	Sequence 355, App
29	145.4	16.0	738	13	US-10-259-165-413	Sequence 413, App
30	145.4	16.0	741	13	US-10-259-165-45	Sequence 45, Appl
31	141.6	15.6	1062	10	US-09-978-730-1	Sequence 1, Appli
32	141.6	15.6	1062	10	US-09-978-729A-1	Sequence 1, Appli
33	141.6	15.6	1062	10	US-09-981-087A-1	Sequence 1, Appli
34	141.6	15.6	1062	10	US-09-978-382A-1	Sequence 1, Appli
35	138.8	15.3	485	13	US-10-259-165-538	Sequence 538, App
36	138.8	15.3	485	13	US-09-853-450-7	Sequence 7, Appli
37	137.6	15.1	735	13	US-10-259-165-713	Sequence 713, App
38	136.8	15.0	665	10	US-09-938-842A-2442	Sequence 2442, Ap
39	136.4	15.0	499	13	US-10-259-165-539	Sequence 539, App
40	136.4	14.8	1098	11	US-09-819-142-15	Sequence 15, Appl
41	134.8	14.8	1098	11	US-09-334-455-357	Sequence 357, App
42	134.8	14.7	747	10	US-09-853-450-27	Sequence 27, Appl
43	133.6	14.7	906	16	US-10-278-536-195	Sequence 195, App
44	133.6	14.7	906	16	US-10-278-536-195	Sequence 2404, Ap
45	133.2	14.7	705	10	US-09-938-842A-2404	

ALIGNMENTS

RESULT 1

US-10-104-580-10
; Sequence 10, Application US/10104580
; Publication No: US2003033628A1
; GENERAL INFORMATION:

; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in poplar and other plant species.
; FILE REFERENCE: 62486
; CURRENT APPLICATION NUMBER: US/10/104,580
; CURRENT FILING DATE: 2002-03-21
; PRIOR FILING DATE: 1999-10-01
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 09/287,700
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: 60/080,851
; PRIOR FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1219
; TYPE: DNA
; ORGANISM: Populus balsamifera subsp. trichocarpa

; NAME/KEY: CDS
; LOCATION: (196)..(921)
US-10-104-580-10

Query Match 28.6%; Score 260; DB 15; Length 1219;
Best Local Similarity 63.7%; Pred. No. 1.4e-65;
Matches 433; Conservative 0; Mismatches 235; Indels 12; Gaps 2;

QY 2 TGGGTCGCGGAAGATTCAGATAAGAGGATTGAAATACCTACCAACGACAGGTCCTT 61

DB 242 TGGGAAGGGGAAAGGTGGAGATCAAGCGGATCGAGAACCAACCATCGCAATGTCCTT 301

QY 62 TCTGCAAGCGCGGAAATGGTTTATTAAAGAGGCGGTATGAATTATCACTTTCTTGTGATG 121

Db 302 TCTGCAAAAGGCGCAGTGGTTTCTCAAGAAAGCCTACGAATTAATCTGTCTTTCGGATG 361
Qy 122 CAGAAGTGGCCCTCATCTGCTTCTCCAGCAGAGGAGACTTATGAAATTTGCCAACACA 181
Db 362 CTGAGGTGACATCATCTGCTTCTAGCGCGGCTGCGCTTATGAGTACTCTAACGATA 421
Qy 182 GCCTGAAGAGGACGATTTGAGAGGTACAAGAAGACTTTCGTTGACAAACACCGAGGGG 241
Db 422 GTCTCAAAATCAACAATTTGAGAGGTACAAGAAGCCTCTCAGATTTCTTCAAAACACTGGGT 481
Qy 242 CGATATCAGAGTCCAAATCTCAGTATTTGGCAACAGAGGCTGTAACTCAGACACAGA 301
Db 482 CTGTTTCTGAAGCAATGCTCAGTACTACCAAGAAAGTGCCTGCAAGCTGCGTTCCCAA 541
Qy 302 TTGACATTTTGCAAAATGCAAAATAGGATTTGATGGGTGACGGGCTTACAGCTTTTGAACA 361
Db 542 TTGGTAATTTGCAGAAATCAACAGGCATATGCTGGGTGAGCGCTTAGTTTCAATGAGTG 601
Qy 362 TTAAGGAATCAAGCAATTTGAGGTTCGACTTGAAGAAAGGAATCAGCCGAGTTCGATCCA 421
Db 602 TGAAGGAATTTAAGAGTTTGGAAATACGACTTCGAGAAAGGAATTAAGCAGAAATTCGTTCCA 661
Qy 422 ABAAGACGAGATGCTTCTTGAAGAGATCGACATCATGACAGAGAGGGAACACATACTTA 481
Db 662 AAAAGATGAGCTGTTGTTTGCAGAAATCGAGTATTCGAGAGAGGGAGGTGACTTGC 721
Qy 482 TCCAGGAGAAATGAGATTTCTTCGAGCAAGATAGCCGAGTGTCA ---GAATAGCCAAACA 538
Db 722 ACAAATTAACCACTTCTCCGAGCAAGATTTTCAAGAGATGAAGAAGCGAGACAGA 781
Qy 539 CGAATCTTATCAGCTCCGGAATATGATGCAC ---GCCCGCATTCGACTCTC 589
Db 782 TGAATTTGATGCCAGGAGGAGCAGCTTTGAGATCGTGCACTTCAACCATATGATCTTC 841
Qy 590 GAAATTTCTACATGCAAAATCTAATCGATGCGGCCCATCACTATGACATTCAGGAACAAA 649
Db 842 GGAATTTCTCAAGTGAATGATGATACAGCTTCAAGCTTCACTACATCAAGATCAGA 901
Qy 650 CAACGCTTCAGCTTGGCTGA 669
Db 902 TGGCCCTTCAGTTAGTTAA 921

RESULT 2
US-10-104-580-11
; Sequence 11, Application US/10104580
; Publication No. US20030033628A1
; GENERAL INFORMATION:
; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; TITLE OF INVENTION: poplar and other plant species.
; FILE REFERENCE: 62486
; CURRENT APPLICATION NUMBER: US/10/104,580
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/410,464
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 09/287,700
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: 60/080,851
; PRIOR FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Populus balsamifera subsp. trichocarpa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(723)
US-10-104-580-11

Query Match 28.6%; Score 259.8; DB 15; Length 723;
Best Local Similarity 63.9%; Pred. No. 1.1e-65;

Matches 431; Conservative 0; Mismatches 232; Indels 12; Gaps 2;
Qy 2 TGGTCTGCGGAGATTTGAGATAAAGAGGATTTGAAATACTACGAAACCGACAGGTCACTT 61
Db 47 TGGAAAGGGGAAGGTGAGATCAAGCGGATCGAAGACACCAATTCGCCAAGTCACTT 106
Qy 62 TCTCAAGCCGCGAAATGCTTTTAAAGAGGCGTATGAATATCAAGTTCTTTTGTGATG 121
Db 107 TCTCAAAAAGCGCAGTGGTTTGTCTAAGAAAGCCTACGAATTTATCTGTCTTTTGGCATG 166
Qy 122 CAGAAGTGGCCCTCATCTCTTCCAGCAGAGGAGACTTTTATGAATTTTGCACCAACA 181
Db 167 CTGAGTTTGCATCTCATCTCTTCTAGCCGCGGTGCGCTTTATGAGTACTCTTAACGATA 226
Qy 182 GCCTGAAGAGACCAATTTGAGAGGTACAAGAGACTTTCGTTGACAAACAACCGAGGGG 241
Db 227 GTGTCAAAATCAACAATTTGAGAGGTACAAGAGGCTCTGCAGATTTCTTCAAACTCGGT 286
Qy 242 CGATATCAGATCCAAATTTCTCAGTATTTGGCAACAGGAGGCTGTAAACTCAGACAAACA 301
Db 287 CTGTTTCTGAAGCAATGCTCAGTACTACCAAGCAAGAGTGCAGAGTGGTTCCTCCAAA 346
Qy 302 TTGACATTTTGCAAAATGCAAAATAGGCAATTTGATGGGTGACGGCTTTACAGCTTTGAACA 361
Db 347 TTGGTAATTTGCAGAAATCAACAGGCATATGCTGGGTGAAGCGCTTAGTTTCAATGAGTG 406
Qy 362 TTAAGGAATCAAGCAATTTGAGGTTCGACTTGAAGAAAGGAATTAAGCAGAAATTCGTTCCA 421
Db 407 TGAAGGAATTTAAGAGTTTGGAAATACGACTTGAAGAAAGGAATTAAGCAGAAATTCGTTCCA 466
Qy 422 AAAAGACGAGATGTTGCTTCAAGAGATCGACATCATGACAGAGAGGGAACACATACTTA 481
Db 467 AAAGAAATGAGCTGTTGTTTGCAGAAATCGAGTATATGCAAGAGAGGAGGTTGACTTGC 526
Qy 482 TCCAGGAAATGAGATTTCTTCGAGCAAGATAGAGCGGAGTGTCA ---GAATAGCCCAACA 538
Db 527 ACAAATTAACCACTTCTCCGAGCAAGATTTTCAAGAGATGAAGAAGCGAGACAGA 586
Qy 539 CGAATCTTATCAGCTCCGGAATATGATGCAC ---GCCCGCATTCGACTCTC 589
Db 587 TGAATTTGATGCCAGGAGGAGCAGACTTTGAGATCGTGCACTCTCAACCATATGACTCTC 646
Qy 590 GAAATTTCTACATGCAAAATCTAATCGATGCGGCCCATCACTATGACATTCAGGAACAAA 649
Db 647 GGAATTTCTCAAGTGAATGGAATTAAGCCTGCAAGCTTCACTACATCAAGATCAGA 706
Qy 650 CAACGCTTCAGCTTGG 664
Db 707 TGGCCCTTCAGTTAG 721

RESULT 3
US-09-978-730-5
; Sequence 5, Application US/09978730
; Patent No. US20020129403A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegen, Sarah
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 19452A-000920US
; CURRENT APPLICATION NUMBER: US/09/978,730
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 959
; TYPE: DNA

```

; LOCATION: (78)..(818)
; OTHER INFORMATION: AGAMOUS-LIKE 5 (AGL5)
US-09-978-729A-5

Query Match      28.1%; Score 255.6; DB 10; Length 959;
Best Local Similarity 68.3%; Pred. No. 2.3e-64;
Matches 354; Conservative 0; Mismatches 164; Indels 0; Gaps 0

Qy 1 ATGGGTCTGGGAGATTGAGATAAAGAGGTTGMAAATACTACGACCGACAGGTCACCT 60
    |||
Db 123 ATAGGAGAGGAGATAGATAAAGAGGATAGAGAACACTACGAATCGTCAAGTCACCT 182
    |||

Qy 61 TTCTGCAAGCGCGAAATGGTTTATTAAAGAAGCGGTATGAATATATCAGTTCTTTCTGTGAT 120
    |||
Db 183 TTCTGCAAGCGCGCAATGGTTTACTCAAGAAAGCTTATGAGTCTCTGTCTTTGTGTGAC 242
    |||

Qy 121 GCAGAGTGGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTATGAATTTGCCAACCA 180
    |||
Db 243 GCTGAGGTTGCTCTTGTCAATCTCTCCACTCGAGGCGGTCTCTACGAGTACGCCAACAA 302
    |||

Qy 181 AGCGTCAAGAGGAGCAGATTGAGAGGTACAAGAAGACTTTGCGTTGCACAAACCAACGAGG 240
    |||
Db 303 AGTGTGAGAGGAACAATAGAAAGGTACAAGAAAGCTTTGCTCCGACGCCGTAAACCCCTCG 362
    |||

Qy 241 GCGATATCAGAGTCCAAATCTCAGTANTTGCACACAGAGGCTGGTAAACTCTCAGACAACAG 300
    |||
Db 363 ACCATACCGAAGCTAACTACGTACTATCAGCAAGAGCGGTCTAAACTCTCGGAGACAG 422
    |||

Qy 301 ATTGCATATTTGCAAAATGCAAAATAGGCATTTGATGGTGCACGGGCTTACAGCTTTGAAC 360
    |||
Db 423 ATTCGGGACATTTCAAAATTTGAAACAGACACATTTCTGGTGAATCTCTTGGTTCCITTGAAC 482
    |||

Qy 361 ATTAAGGAACTCAAGCACTTGAGGTTTCGACTTTGMAAAAGGAATCAGCCGAGTGGCATCC 420
    |||
Db 483 TTTTAAGAACTCAAGAACCTTGAAGTAGGTTGAGAAAGGAATCAGTGGTCCGATCC 542
    |||

Qy 421 AAAAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATGCGAGAGAGGGGAACAACATACCT 480
    |||
Db 543 AAGAAGCAGCAGATGTTAGTTGCGAGATTGAATACATGCAAAAAAGGGAAATTCGAGCTG 602
    |||

Qy 481 ATCCAGGAGATGAGATTCTTCGCGACGACAGATGCCGA 518
    |||
Db 603 CAAAACGATACATGTTATCTCCGCTCCAGATTACTGA 640
    |||

```

```

RESULT 5
US-09-981-087A-5
; Sequence 5, Application US/09981087A
; Patent No. US20020178466A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegren, Sarah
; APPLICANT: Farrandiz, Cristina
; APPLICANT: the Regents of the University of California
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 19452A-000940US
; CURRENT APPLICATION NUMBER: US/09/981,087A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(818)
; OTHER INFORMATION: AGAMOUS-LIKE 5 (AGL5)

```

```

; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(818)
; OTHER INFORMATION: AGAMOUS-LIKE 5 (AGL5)
US-09-978-730-5

Query Match      28.1%; Score 255.6; DB 10; Length 959;
Best Local Similarity 68.3%; Pred. No. 2.3e-64;
Matches 354; Conservative 0; Mismatches 164; Indels 0; Gaps 0

QY 1 ATGGGTGCGGGAAGATTGAGATAAAGAGGATTGAAATACTACGAACCGACAGGTCAC 60
Db 123 ATAGGGAGAGGGAAGATAGATAAAGAGGATAGAGAACACTACGAATCGTCAAGTCAC 182
QY 61 TTCTGCAAGCGCCGAAATGGTTATTATTAAGAGAGCGGTATGAATTATCAGGTTCTTCTGAT 120
Db 183 TTCTGCAACGACGCAATGGTTTACTCAAGAAAGCTTATGAGCTCTCTGTTGTGTGAC 242
QY 121 GCAGAAGTGGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTATGAATTTGCCAACCA 180
Db 243 GCTGAGGTTGCTCTTGTGCATCTTCTCCACTCGAGGCCGCTCTACGAGTACGCCAACAA 302
QY 181 ACSGTGAGAGAGACGATTTGAGAGGTACAAAGAGACTTGCCTTGCAACAACCAACGAGG 240
Db 303 AGTGTGAGAGGAAACATTAAGAGGTACAAGAAAGCTTGCTCCGACGCCGTTAACCCCTCG 362
QY 241 GCGATATCAGAGTCCAATCTCAGPATTTGGCAACAGGAGGCTGGTAAACTCAGACAACAG 300
Db 363 ACCATCACCGAAGCTAATACTCAGTACTATCAGCAAGAGGCGTCTAACTCCGGAGACAG 422
QY 301 ATTGACATTTTGCACAAATGCAAAATAGGCGATTTGATGGTGCAGCGGCTTACAGCTTTGA 360
Db 423 ATTGGGACATTCAGAAATTTGNAACGACACATTTCTGGTGAATCTCTTGGTTCCTTGAAC 482
QY 361 ATTAAAGGAATCAAGCAACTTGAGGTTTCGACTTGAAAGGAATACGCGAGTGCAGATCC 420
Db 483 TTTAAAGGAATCAAGAACCTTGAAGTAGGCTTGAGAAAGGAATCAGTCGTGTGCCGATCC 542
QY 421 AAAAAAGAACGAGATGTTGCTTGAAGAGATCCACATCATGCAGAGAAGGGAACAACATAC 480
Db 543 AAGAGCACGAGATGTTAGTTCGAGAGATTGATATCATGCAAAAAGGGAATTCGAGCTG 602
QY 481 ATCCAGGGAATGAGATTTCTTCGACGACGAAGATGCCGA 518
Db 603 CAAAACGATAACATGTATCTCGGTCCCAAGATTACTGA 640

```

```

RESULT 4
US-09-978-729A-5
; Sequence 5, Application US/09378729A
; Patent No. US20020178465A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljgren, Sarah
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 19452A-0009500S
; CURRENT APPLICATION NUMBER: US/09/978,729A
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS

```

US-09-981-087A-5

Query Match	28.1%;	Score 255.6;	DB 10;	Length 959;
Best Local Similarity	68.3%;	Fred. No. 2.3e-04;	Indels 0;	Gaps 0;
Matches 354;	Conservative 0;	Mismatches 164;		
QY	1	ATGGGTCGTGGGAAGATTGAGATAAGAGGATTGAAATACTACGAACCGACAGTGCCTACT	60	
Db	123	ATAGGGAGAGGGAGATAGATAGATAAAGAGGATAGAGAACACTACGAATCGTCAAGTCTACT	182	
QY	61	TTCTCCAGAGCCCGGAATGGTTTATTAAAGAAAGCGGTATGAATTTATCAGTCTCTTTGTGTGAT	120	
Db	183	TTCTTCGAACAGCAGCAATGGTTTACTCTAGAAAGCTTATGAGTCTCTGTCTTGTGTGAC	242	
QY	121	GCAGAAGTGGCCCTCATCGTCTTCTCCAGCAGAGGGAGACTTTATGAATTTGCCAACCAAC	180	
Db	243	GCTGAGGTGCTCTTCTGTCATCTTCTCCACTCGAGGCCGTCTCTACGAGTACGCCAACAC	302	
QY	181	ACGGTGNAGAGGACGATTGAGAGGTACAGAAGACTTTGCCTTGCACAAACACACGAGGG	240	
Db	303	AGTGTGAGAGGAACAATAGAAAGGTACAAAGAGCTTGTCCGAGCGCTTATACCTCCG	362	
QY	241	GGGATATCAGAGTCCCAATTCTCAGTATTGGCAACAGAGGCTCGTTAAACTCAGACACACAG	300	
Db	363	ACCATCCCGAAGCTAACTACTCAGTACTATCAGCAAGAGGCGTCTAAACTCCGGAGACAG	422	
QY	301	ATTGACATTTTCRAAATGCCAAATAGGCAATTTGATGGTGCACGGGCTTACAGCTTTGAC	360	
Db	423	ATTCCGGCAATTCAGAAATTTGAACAGACACATCTTGTGTGAATCTCTTGGTCCITTGAC	482	
QY	361	ATTAAAGGAATCAAGCAACTTGAGGTTTCGACTTGCAAAAGGAATTCAGCCGAGTGCAGTCC	420	
Db	483	TTTAAAGGAATCAAGAACTTGAAAGTAGTAGGCTTGAGAAAGGAATCAGTGTGTGCCGATCC	542	
QY	421	AAAAGAAACGAGATCTTCTTGAGAGATCCACATCATGCGAGAGAGGGGAACACATACTTT	480	
Db	543	AGAAGCACGAGATGTTAGTTCGAGAGATTGAATACATGCAAAAAGGGAAATCGAGCTG	602	
QY	481	ATCCGAGGAATGAGATTTCTTCGACGCAAGATAGCCGA	518	
Db	603	CAAAACGATAACATGTATCTCCGCTCCCAAGATTACTGA	640	

RESULT 6

```

US-09-978-382A-5
; Sequence 5, Application US/09978382A
; Publication NO. US20020194647A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegen, Sarah
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; TITLE OF INVENTION: Plants
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 19452A-000930US
; CURRENT APPLICATION NUMBER: US/09/978,382A
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(818)
; OTHER INFORMATION: AGAMOUS-LIKE 5 (AGL5)
US-09-978-382A-5

```

Query Match

28.1%; Score 255.6; DB 10; Length 959;

	Best Local Similarity	58.3%;	Pred. No. 2.3e-64;		0;
	Matches	354;	Conservative	0;	Mismatches 164;
					Indels 0;
					Gaps 0;
QY	1	ATGGGTCTGGGAAGATTGAGATAAAGAGAGATTGAAAACTACTAGAACCGACAGGTCACT	60		
Db	123	ATAGGAGAGGGGAAGATAGAGATAAAGAGGATAGAGAACACTACGAATCTCGTCAAGTCAC	182		
QY	61	TTCTGCAGCGCCGAATATGGTTTATTAAAGACGGCTATCAAAATATCAGTTCCTTTGFGAT	120		
Db	183	TTCTGCAAAACGACGCAATGGTTTACTCAAGAAAGCTTATAGAGTCTCTGTCTTGTGTGAC	242		
QY	121	GCAGAAAGTGCCCTCATCTGTTCTCCAGCAGAGGGAGACTTTTATGAATTTGCCAACCA	180		
Db	243	GCTGAGGTTGCTCTTGTCTCATCTTCTCCACTCGAGGCGCTCTTACGAGTACGCCAACAA	302		
QY	181	AGCGTGAAGAGAGCGATTGAGAGGTTACAGAAAGACTTTCGTTGACAACACACCGAGGG	240		
Db	303	AGTGTGAGAGGAACAATAGAAAAGGTCAAGAAAGCTTGTCTCCAGCGCGTTAACCCCTCG	362		
QY	241	CGCATATCAGAGTCCAAATCTCTAGTATTGGCAACAGGAGGCTGGTAAACTCAGACACAG	300		
Db	363	ACCATCACCAGAGTAATATCTAGTACTATCAGCAGAGGGGCTAAACTCCGAGACAG	422		
QY	301	ATTGACATTTTGCAAAATGCAAAATAGSCATTTGATGGTGACGGGCTTACAGCTTTGAAC	360		
Db	423	ATTCCGGACATTGAGAAATTTGAACAGACACATTTCTTGGTGAATCTCTTGGTTCCTTGAAC	482		
QY	361	ATTAAGAACTCAAGCAACTTGAAGTTGAGTTGCACTTGAANAAGGAATCAGCCGAGTCGGATCC	420		
Db	483	TTTAAAGAACTCAAGAACCTTTGAAGTAGGCTTTAGAAAGGAATTCAGTCTGTGTCCGATCC	542		
QY	421	AAAAGAAACAGAGATGTTTGGTTTGAAGAGATCGACATCATGCAAGAGGGAACACATATCTT	480		
Db	543	AAGNAGCAGAGATGTTAGTTTGCAGAGATTTGAATATCATGCAAAAAGGGAAATCGAGCTG	602		
QY	481	ATCCAGAGAAATGAGATCTCTTCGCAGCAAGATAGCCGA	518		
Db	603	CAAAACCATTAACATGTATCTCCGCTCCAAAGATTACTGA	640		

RESULT 7

```

US-09-978-740A-5
; Sequence 5, Application US/09978740A
; Publication NO. US2003005481A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegren, Sarah
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; TITLE OF INVENTION: Plants
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 19452A-000960US
; CURRENT APPLICATION NUMBER: US/09/978,740A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 959
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(818)
; OTHER INFORMATION: AGAMOUS-LIKE 5 (AGL5)
; US-09-978-740A-5

```

Query Match 28.1%: Score 255.6: DB 11: Length 959;

Query Match	28.1%;	Score 255.6;	DB 11;	Length 959;
Best Local Similarity	68.3%;	Pred. No. 2.3e-64;		
Matches 354; Conservative	0;	Mismatches 164;	Indels 0;	Gaps 0;

QY 1 ATGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAAAATCTACGAACGACAGGTCACCT 60
Db 123 ATAGGGAGAGGAGATAGAGATTAAGAGATAGAGAACTACGATCGTCAAGTCACCT 182
QY 61 TTCTGCAAGCGCCGAATGGTTTATTAAAGAGCGGTATGAATATCAGTTCCTTTGTGAT 120
Db 183 TTCTGCAAAACGACGCAATGGTTTACTCAAGAAAGCTTATGAGCTCTCTGTCTTGTGTGAC 242
QY 121 GCAGAACTGGCCCTCATCGTCTCTCCAGCAGAGGGAGACTTTATGAATTTGCGCAACAC 180
Db 243 GCTGAGGTTGCTCTGTGCACTCTCTCCATCTCGAGGCGCTCTCAGAGTACGCCAACAC 302
QY 181 AGCGTGAAGAGGAGATGAGAGGTACAAGAGACTTTCGTTGACAAACACACGAGGG 240
Db 303 AGTGTGAGAGGAAACAATAGAAAGGTACAAGAAAGCTTGTCCGACGCGCTTAAACCCCTCCG 362
QY 241 GCGATATCAGAGTCCCAATCTCAGTATTGSCACAGGAGGCTGGTAAACTCAGACACACAG 300
Db 363 ACCATCAGGAGCTAATCTCAGTACTATCAGCAGAGGCGGTCTAACTCCGAGACAG 422
QY 301 ATTGACATTTTGCAAAATGCAATPAGCATTTTGATGGGTGACGCGGTTCACAGCTTTGAAC 360
Db 423 ATTCGGGACATTCAGATTTGAACAGACACATCTCTGGTGAATCTCTTGGTTCCTTGAAC 482
QY 361 ATTAAGGAACCTCAAGCACTTGGGTTGAGTTGCACTTGAAGAAAGAAATCAGCCGAGTCCATCC 420
Db 483 TTTAAGGAACCTCAAGAACCTTGAAGTAGGCTTGAAGAAAGAAATCAGTCTGTCGATCC 542
QY 421 AAAAAGAACAGAGATGTTGCTTGAAGAGATCGACATCATGACAGAGAGGGAACACATACCT 480
Db 543 AAGAGACAGAGATGTTAGTTGACAGATTTGAATACATGCAAAAGGGAATCGAGCTG 602
QY 481 ATCCAGGAGATGAGATCTTCGCGCAAGATAGCCGA 518
Db 603 CAAAACGATACATGATCTATCCGCTCCAAGATTACTGA 640

RESULT 8
US-10-104-580-15
; Sequence 15, Application US/10104580
; Publication No. US20030033628A1
; GENERAL INFORMATION:
; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; TITLE OF INVENTION: poplar and other plant species.
; FILE REFERENCE: 62486
; CURRENT APPLICATION NUMBER: US/10/104,580
; PRIOR APPLICATION NUMBER: 2002-03-21
; PRIOR FILING DATE: 2002-03-21
; PRIOR FILING DATE: 09/410,464
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 09/287,700
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: 60/080,851
; PRIOR FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 714
; TYPE: DNA
; ORGANISM: Populus balsamifera subsp. trichocarpa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(714)
US-10-104-580-15
Query Match 27.3%; Score 248.2; DB 15; Length 714;
Best Local Similarity 67.0%; Pred. No. 2.9e-62;
Matches 352; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 2 TGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAAAATCTACGAACCGACAGGTCACCT 61
Db 47 TGGGAGGGGAAAGGTGAGATCAAGCGATCGAGACACACCAATCGCCAGTCACTT 106

QY 62 TCTCAAGCGCCGAATGGTTTATTAAAGAGCGGTATGAATATCAGTTCCTTTGTGATG 121
Db 107 TCTCAAAAGCGGNAUGTTTGTCTCAAGAAAGCTATGAATATCTGTTCTTTGGCGATG 166
QY 122 CAGAAAGTGGCCCTCATCGTCTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACACACA 181
Db 167 CTGAGGTTGCACTCATCGTCTCTCCAGCGCGCTTATGAGTACTCTTAACAATA 226
QY 182 GCGTGAAGAGACAGATTGAGAGGTACAAAGAGACTTGGGTTGACAAACCAACCGAGGGG 241
Db 227 GTGTCAATCTCAATTTGAAGGTACAAAGGCAATGTGCAGATTTCTTCAACAACCGGT 286
QY 242 CCAATATCAGAGTCCAAATCTCAGTATTGGCAACAGAGGCTGGTAAACTCAGACACACA 301
Db 287 CAGTTTCTGAAGCCCAATCTCAGTCTCTATCAGCAAGAGCTGCCAAGCTCGCTCGCAA 346
QY 302 TTGACATTTTGCMAAATGCAATAGGCAATTTGATGGGTGACGCGGTTACAGCTTTGAACA 361
Db 347 TTGGTAAATTTGCAGAAATTCAAACAGGATATGCTGGGTGAATCACTTAGTGCATTGAGTG 406
QY 362 TTAAGGAACCTCAAGCAACTTGGGTTGCACTTTGAAAAGGAATCAGCCGAGTGGGATCCA 421
Db 407 TGAAGGAACCTTAAGAGCTTGGAGATTAATACTTGAGAAAGGAATTTGGTGAATTCGTTCCA 466
QY 422 AAAAAGAACAGAGATGTTGCTTGAAGAGATCGACATCATGACAGAGAGGGAACACATACCTA 481
Db 467 AAAAGAAATGAGCTGTGTTGCTGAAATGAGTATATGCAAGAGGAGGATTTGACTTGC 526
QY 482 TCCAGGAGAAATGAGATTTCTTCGAGCAAGATAGCCGAGTGTGAGA 526
Db 527 ACAACAATAACCGAGCTTCTCCGAGCAAGATTGCAGAGAAATGAAA 571

RESULT 9
US-10-104-580-14
; Sequence 14, Application US/10104580
; Publication No. US20030033628A1
; GENERAL INFORMATION:
; APPLICANT: Strauss et al.
; TITLE OF INVENTION: Floral homeotic genes for manipulation of flowering in
; TITLE OF INVENTION: poplar and other plant species.
; FILE REFERENCE: 62486
; CURRENT APPLICATION NUMBER: US/10/104,580
; PRIOR APPLICATION NUMBER: 2002-03-21
; PRIOR FILING DATE: 09/410,464
; PRIOR FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 09/287,700
; PRIOR FILING DATE: 1999-04-06
; PRIOR APPLICATION NUMBER: 60/080,851
; PRIOR FILING DATE: 1998-04-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 1159
; TYPE: DNA
; ORGANISM: Populus balsamifera subsp. trichocarpa
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (99)..(815)
US-10-104-580-14
Query Match 27.3%; Score 248.2; DB 15; Length 1159;
Best Local Similarity 67.0%; Pred. No. 3.9e-62;
Matches 352; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 2 TGGGTCGTGGGAAGATTGAGATAAAGAGGATTGAAAAATCTACGAACCGACAGGTCACCT 61
Db 145 TGGGAGGGGAAAGGTGAGATCAAGCGATCGAGACACACCAATCGCCAGTCACTT 204
QY 62 TCTGCAAGCGCCGAATGGTTTATTAAAGAGCGGTATGAATATCAGTTCCTTTGTGATG 121
Db 205 TCTGCAAAAGCGGGAATGGTTGCTCAAGAAAGCGCTATGAATATCTGTTCTTTGGCGATG 264

QY 122 CAGAGTGGCCCTCATCTCTCTCCAGCAGAGGAGACTTTATGATTTGCCAACACA 181
Db 265 CTGAGGTTGCATCTATCTCTCTCCAGCAGGAGGAGCTTTATGAGTACTCTAACAA 324
QY 182 GCGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241
Db 325 GTGTCAAACTCAATTTGAAGGAGTACAAAGGAGGAGGAGGAGGAGGAGGAGGAG 384
QY 242 CGATATCAGATCCCAATCTCAGTATTCGCAACAGGAGGAGGAGGAGGAGGAGGAG 301
Db 385 CAGTTTCTGAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 444
QY 302 TTGACATTTTCAAAATGCAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361
Db 445 TTGGTAATTTGAGAAATCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 504
QY 362 TTAAGGAATCAAGCAATCTGAGTATTCGCAACAGGAGGAGGAGGAGGAGGAGGAG 421
Db 505 TGAAGGAATCAAGCAATCTGAGTATTCGCAACAGGAGGAGGAGGAGGAGGAGGAG 564
QY 422 AAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATCGACAGAGGAGGAGGAGG 481
Db 565 AAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATCGACAGAGGAGGAGGAGG 624
QY 482 TCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 526
Db 625 ACAACAAATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 669

RESULT 10
US-09-978-730-3
; Sequence 3, Application US/09978730
; Patent No. US20020129403A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegen, Sarah
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; FILE REFERENCE: 19452A-000920US
; CURRENT APPLICATION NUMBER: US/09/978,730
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; NAME/KEY: CDS
; LOCATION: (7)..(753)
; OTHER INFORMATION: AGAMOUS-LIKE 1 (AGL1)
US-09-978-730-3

Query Match 27.0%; Score 245; DB 10; Length 896;
Best Local Similarity 67.1%; Pred. No. 2.9e-61;
Matches 347; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 2 TGGGTCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 61
Db 53 TAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 112
QY 62 TCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 121
Db 113 TCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 172
QY 122 CAGAGTGGCCCTCATCTCTCTCCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 181
Db 173 CCGAAGTTGGCCCTCATCTCTCTCCAGCAGAGGAGGAGGAGGAGGAGGAGGAGG 232

QY 182 GCGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241
Db 233 GTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 292
QY 242 CGATATCAGATCCCAATCTCAGTATTCGCAACAGGAGGAGGAGGAGGAGGAGGAG 301
Db 293 CCGTCAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 352
QY 302 TTGACATTTTCAAAATGCAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361
Db 353 TCCAGATATTCAGAAATCAATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 412
QY 362 TTAAGGAATCAAGCAATCTGAGTATTCGCAACAGGAGGAGGAGGAGGAGGAGGAG 421
Db 413 TCAAGGAATCAAGCAATCTGAGTATTCGCAACAGGAGGAGGAGGAGGAGGAGGAG 472
QY 422 AAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATCGACAGAGGAGGAGGAGG 481
Db 473 AAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATCGACAGAGGAGGAGGAGG 532
QY 482 TCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 518
Db 533 AACACATATCATCTCTCTCCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 569

RESULT 11
US-09-978-729A-3
; Sequence 3, Application US/09978729A
; Patent No. US20020178465A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegen, Sarah
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; FILE REFERENCE: 19452A-000950US
; CURRENT APPLICATION NUMBER: US/09/978,729A
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; NAME/KEY: CDS
; LOCATION: (7)..(753)
; OTHER INFORMATION: AGAMOUS-LIKE 1 (AGL1)
US-09-978-729A-3

Query Match 27.0%; Score 245; DB 10; Length 896;
Best Local Similarity 67.1%; Pred. No. 2.9e-61;
Matches 347; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 2 TGGGTCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 61
Db 53 TAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 112
QY 62 TCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 121
Db 113 TCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 172
QY 122 CAGAGTGGCCCTCATCTCTCTCCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 181
Db 173 CCGAAGTTGGCCCTCATCTCTCTCCAGCAGAGGAGGAGGAGGAGGAGGAGGAGG 232
QY 182 GCGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241

Db 233 GTGTGAGGGGTCAATTTGAAAGGTACAAAGAAAGCTTTGTCGATGCGGTCAACCCCTCCTT 292
Qy 242 CGATATCAGAGTCCAAATTTCTCAGTATTTGGCAACAGAGGAGCTGTGTAATCTCAGACAAACAGA 301
Db 293 CCGTCACCGAAGCTAATACTCAGTACTATCAGCAAGAAGCCTCTAAGCTTCGGAGGCAGA 352
Qy 302 TTGACATTTTGC AAAATGCAATAGGCAATTTGATGGGTGACGGCTTACAGCTTTGAACA 361
Db 353 TTCGAGATATTCAGAAATTCAAATAGGCATATTTGTTGGGAATCACATTGGTTCCTTGAAC 412
Qy 362 TTAAGGAATCAAGCAACTTGGAGTTCGATTTGAAAAAGGAATCAGCCGAGTGCAGTCCA 421
Db 413 TCAAGGAATCAAAAACCTAGAGGAGCTTTTGA AAAAGGAATCAGCCGCTGTCGCTCCA 472
Qy 422 AAAAGACGAGATGTTGTTGAAGAGATCGACATCATCAGAGAGAGGAAACACATACCTTA 481
Db 473 AAAAGATGAGCTGTTAGTGGCAGAGATAGATATATGCAAGAGAGGAAATGGAAGTTC 532
Qy 482 TCCAGGAGATGAGATCTTCGCGCAGCAGATAGCCGA 518
Db 533 AACCAATACATGCTCGGAGCAAGATAGCCGA 569

RESULT 12

US-09-981-087A-3
; Sequence 3, Application US/09981087A
; Patent No. US20020178466A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegen, Sarah
; APPLICANT: Farrandiz, Cristina
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 19452A-000940US
; CURRENT APPLICATION NUMBER: US/09/981,087A
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)...(753)
; OTHER INFORMATION: AGAMOUS-LIKE 1 (AGL1)
US-09-981-087A-3

Query Match 27.0%; Score 245; DB 10; Length 896;

Best Local Similarity 67.1%; Pred. No. 2.9e-61;

Matches 347; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGGTCTGTGGAGATTGAGATAAGAGGATTGAAAAATCTACGAACCGACAGTCACTT 61
Db 53 TAGGAGAGGCAAAATAGAGATAAGAGGATAGAGAACACAAATCGTCAAGTTACTT 112
Qy 62 TCTGCAAGCCGAAATGGTTTATAAGAGGCGGTATGATATCATGTTCTTTGTGATG 121
Db 113 TCTGCAACAGCAGCAATGGTCTTCTCAAGAAAGCTTATGAACCTCTCTGTTGTGATG 172
Qy 122 CAGAAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACACA 181
Db 173 CCGAAGTTGCCCTCGTCATCTTCTCCACTCGTGGCCGCTCTATGAGTAGCCCAACA 232
Qy 182 CGGTGAGAGGACGATTGAGAGGTACAAGAGACTTGGGTGACAAACCAACGAGGGG 241
Db 233 GTGTGAGGGGTACAAATTTGAAAGGTACAGAAAGCTTGTTCGATGCGGTCAACCCCTCCTT 292

Qy 242 CGATATCAGAGTCCAAATTTCTCAGTATTTGGCAACAGGAGGCTGTGTAATCTCAGACAAACAGA 301
Db 293 CCGTCACCGAAGCTAATACTCAGTACTATCAGCAAGAAGCCTCTAAGCTTCGGAGGCAGA 352
Qy 302 TTGACATTTTGC AAAATGCAATAGGCAATTTGATGGGTGACGGCTTACAGCTTTGAACA 361
Db 353 TTCGAGATATTCAGAAATTCAAATAGGCATATTTGTTGGGAATCACATTGGTTCCTTGAAC 412
Qy 362 TTAAGGAATCAAGCAACTTGGAGTTCGATTTGAAAAAGGAATCAGCCGAGTGCAGTCCA 421
Db 413 TCAAGGAATCAAAAACCTAGAGGAGCTTTTGA AAAAGGAATCAGCCGCTGTCGCTCCA 472
Qy 422 AAAAGACGAGATGTTGTTGAAGAGATCGACATCATCAGAGAGAGGAAACACATACCTTA 481
Db 473 AAAAGATGAGCTGTTAGTGGCAGAGATAGATATATGCAAGAGAGGAAATGGAAGTTC 532
Qy 482 TCCAGGAGATGAGATCTTCGCGCAGCAGATAGCCGA 518
Db 533 AACCAATACATGCTCGGAGCAAGATAGCCGA 569

RESULT 13

US-09-978-382A-3
; Sequence 3, Application US/09978382A
; Publication No. US20020194647A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegen, Sarah
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 19452A-000930US
; CURRENT APPLICATION NUMBER: US/09/978,382A
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 896
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)...(753)
; OTHER INFORMATION: AGAMOUS-LIKE 1 (AGL1)
US-09-978-382A-3

Query Match 27.0%; Score 245; DB 10; Length 896;

Best Local Similarity 67.1%; Pred. No. 2.9e-61;

Matches 347; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 2 TGGGTCTGTGGAGATTGAGATAAGAGGATTGAAAAATCTACGAACCGACAGTCACTT 61
Db 53 TAGGAGAGGCAAAATAGAGATAAGAGGATAGAGAACACAAATCGTCAAGTTACTT 112
Qy 62 TCTGCAAGCCGCAATGGTTTATAAGAGGCGGTATGATATCATGTTCTTTGTGATG 121
Db 113 TCTGCAACAGCAGCAATGGTCTTCTCAAGAAAGCTTATGAACCTCTCTGTTGTGATG 172
Qy 122 CAGAAGTGGCCCTCATCGTCTTCTCCAGCAGAGGAGACTTTATGAATTTGCCAACACA 181
Db 173 CCGAAGTTGCCCTCGTCATCTTCTCCACTCGTGGCCGCTCTATGAGTAGCCCAACA 232
Qy 182 CGGTGAGAGGACGATTGAGAGGTACAAGAGACTTGGGTGACAAACCAACGAGGGG 241
Db 233 GTGTGAGGGGTACAAATTTGAAAGGTACAGAAAGCTTGTTCGATGCGGTCAACCCCTCCTT 292
Qy 242 CGATATCAGAGTCCAAATTTCTCAGTATTTGGCAACAGAGGCTGTGTAATCTCAGACAAACAGA 301
Db 293 CCGTCACCGAAGCTAATACTCAGTACTATCAGCAAGAAGCCTCTAAGCTTCGAGGCAGA 352

QY	301	ATTGACATTTTGCATAATGCAATAGGCATTTGATGGTGACGGGCTTACAGCTTTGAAC	360
Db	320	ATCCAAACGATTCARAACTCCACAGGAATCTGATGGGAGACTCTTTGAGTTCTTAACT	379
QY	361	ATTAAGGAATCTAAGCAACTTGAGGTTGCACTTGAAAGGAATCAGCCGAGTGCATCC	420
Db	380	GTCAAGGAATCAAAACAAGTTGAGAAATCGGCTTGAGAAAGCTATCTTAGGATCAGGTCC	439
QY	421	AAAAAGAACGAGATGTTGCTTGAAGAGATCGACATCATGCAGAGAAAGGGAACACATACTT	480
Db	440	AAGAGCATGAGTTGCTTCTAGTTGAATCGAAACCGCGAAGGAGGAGATTGAGCTT	499
QY	481	ATCCAGGAGAAATGAGATTTCTTCGAGCAAGATAGCCG-AGTGTCAAGATAGCCCAACAC	539
Db	500	GACAAATCAGAACATCTATCTAAGAACTAAGGTAGCAGAGTGGAGAGGTATCAACACAC	559
QY	540	GACATGTTAT	550
Db	560	CATCATCAAT	570

Search completed: January 22, 2004, 01:20:39
Job time : 310.264 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 17:02:36 ; Search time 2018.55 Seconds
(without alignments)
12185.050 Million cell updates/sec

Title: US-09-936-869-5
Perfect score: 1012
Sequence: 1 acgaattaagtcgaatggc.....ttctatggtataaaaaaa 1012

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 4552784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```
EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_man.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_rod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*
29: gb_gss2.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	634	62.6	635	9 AV825172	AV825172 AV825172
2	521	51.5	530	9 AI995637	AI995637 701676626
3	520	51.4	520	9 AV527174	AV527174 AV527174
4	491	48.5	607	9 AV832175	AV832175 AV832175

```

C 5 490 48.4 490 9 AB038725
C 6 457.6 45.2 527 9 AV520789
C 7 420.4 41.5 422 9 AV788022
C 8 417.4 41.2 429 9 AV820692
C 9 416.4 41.1 418 9 AV819666
C 10 402 39.7 418 9 AV819464
C 11 391.4 38.7 394 14 Z26559
C 12 378.8 37.4 570 13 BQ791108
C 13 375.6 37.1 400 9 AV816618
C 14 371 36.7 371 9 AV527256
C 15 365.2 36.1 453 14 R90379
C 16 359.4 35.5 361 9 AV532040
C 17 349.8 34.6 849 12 BM359116
C 18 343 33.9 351 10 BE524429
C 19 339.4 33.5 684 9 AI72386
C 20 331.2 32.7 771 14 CB292632
C 21 328 32.4 685 10 BG585152
C 22 319.6 31.6 865 14 CB292631
C 23 315.4 31.2 941 9 AI729649
C 24 311.6 30.8 630 13 BQ405390
C 25 306.4 30.3 690 9 AW039494
C 26 306.2 30.3 664 14 CA826107
C 27 305 30.1 627 13 BU825514
C 28 304 30.0 804 14 CA801284
C 29 303.2 30.0 392 14 T23025
C 30 302 29.8 741 14 CA799030
C 31 301.4 29.8 632 10 BE191140
C 32 301.4 29.8 646 9 AW775880
C 33 294.4 29.1 645 10 BG597684
C 34 290.2 28.7 658 10 BG455225
C 35 285.4 28.2 351 14 T44541
C 36 285.4 28.2 594 12 B1420700
C 37 285.4 28.2 674 10 BG581199
C 38 280 27.7 646 10 BE124916
C 39 279.2 27.6 651 13 BQ119665
C 40 279 27.5 567 10 BF113095
C 41 278.4 27.5 590 10 BE435686
C 42 274.6 27.5 613 10 BE020073
C 43 274.6 27.1 670 10 AW980478
C 44 274.4 27.1 702 9 AW223027
C 45 271.4 26.8 554 9 AW775566

```

ALIGNMENTS

```

RESULT 1
AV825172
LOCUS AV825172 635 bp mRNA linear EST 01-APR-2002
DEFINITION AV825172 RAF16 Arabidopsis thaliana cDNA clone RAFL06-88-B05 5',
mRNA sequence.
ACCESSION AV825172
VERSION AV825172.1 GI:19867232
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 635)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satoh, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.,
and Shinozaki, K.
REFERENCE
AUTHORS
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp

```

An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

Source Location/Qualifiers
 1..635
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAFL06-88-B05"
 /dev stage="plants at various developmental stages from germination to mature seeds"
 /lab host="DH10B"
 /clone lib="RAFL6"
 /note="Site 1: SstI; Site 2: XhoI; subjected to dehydration (1, 2, 5, 10, 24 hr) treatments"
 154 a 138 c 145 g 198 t

BASE COUNT

ORIGIN
 Query Match 62.6%; Score 634; DB 9; Length 635;
 Best Local Similarity 100.0%; Pred. No. 1.4e-135;
 Matches 634; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AATGGCGTCACGTTATGTTCTTCCTTCGTTGCGTGTATCGCGGAGCATTTGCCGG 74
 Db 2 AATGGCGTCACGTTATGTTCTTCCTTCGTTGCGTGTATCGCGGAGCATTTGCCGG 61
 QY 75 AGAGTTCATCGAATCAATCGATCTAGAGGAGTTCGATTATTCGCTATCTCTTCA 134
 Db 62 AGAGTTCATCGAATCAATCGATCTAGAGGAGTTCGATTATTCGCTATCTCTTCA 121
 QY 135 ATGGCTCGGAACCTATTGCGGTGAATCGCCATGTTGTTCGCAAAACGCTTGTGCGAG 194
 Db 122 ATGGCTCGGAACCTATTGCGGTGAATCGCCATGTTGTTCGCAAAACGCTTGTGCGAG 181
 QY 195 AGGCTCGGATCTCCAACTCAATTCATCGGTTATGCGCTGACATATAACGATGG 254
 Db 182 AGGCTCGGATCTCCAACTCAATTCATCGGTTATGCGCTGACATATAACGATGG 241
 QY 255 TTCGTGGCGCTTCATGTTGTTATCGATCTGATCTTAAGAGAGAGATTTCAACGTTGAT 314
 Db 242 TTCGTGGCGCTTCATGTTGTTATCGATCTGATCTTAAGAGAGAGATTTCAACGTTGAT 301
 QY 315 GGATGGTCTTGAGAGTACTGGCTTAGTCTCAGTTGGTTCTCCATCATCGCATGG 374
 Db 302 GGATGGTCTTGAGAGTACTGGCTTAGTCTCAGTTGGTTCTCCATCATCGCATGG 361
 QY 375 TGGGAAGGTCATTTTGGGCGCCAGAGTGGGAGAACATGGGACTTGTCTTCTCCTGT 434
 Db 362 TGGGAAGGTCATTTTGGGCGCCAGAGTGGGAGAACATGGGACTTGTCTTCTCCTGT 421
 QY 435 TTTTCATGATGAGTATAATTACTTCTTACACACTTAATCTCTTACTTGAAGCATATGT 494
 Db 422 TTTTCATGATGAGTATAATTACTTCTTACACACTTAATCTCTTACTTGAAGCATATGT 481
 QY 495 CACGATGCTCTTTATCAAGCTGCTATGTTGCTTCCAAAGTGAAGATATCTCTAGG 554
 Db 482 CACGATGCTCTTTATCAAGCTGCTATGTTGCTTCCAAAGTGAAGATATCTCTAGG 541
 QY 555 AGGTATCGTAAACAGCATTCAGATGCAATTCATATCACCCCTCGAAGTGGTTTCCAAAG 614
 Db 542 AGGTATCGTAAACAGCATTCAGATGCAATTCATATCACCCCTCGAAGTGGTTTCCAAAG 601
 QY 615 AGATCGAATCGATGAATACGATATGCTTCTAT 648
 Db 602 AGATCGAATCGATGAATACGATATGCTTCTAT 635

RESULT 2
 A1995637/c
 LOCUS

530 bp mRNA linear EST 08-SEP-1999

DEFINITION

701676626 A. thaliana, Columbia Col-0, inflorescence-1 Arabidopsis thaliana cDNA clone 701676626, mRNA sequence.

ACCESSION

A1995637

VERSION

A1995637.1 GI:5842542

KEYWORDS

EST.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

REFERENCE

Chen, J., Moriyan, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Gueffier, K., Kim, C., Doyle, M., Brzozka, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kaestury, K., Borillo, C., Carpio, I., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrega, A., Murry, L., Turner, C., Kikorian, S., Elder, L., and Hanson, D.

AUTHORS

Unpublished

TITLE

Arabidopsis thaliana Gene Expression MicroArray

JOURNAL

Unpublished

COMMENT

Contact: David Smoller, Ph.D.
 Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
 4633 World Parkway Circle, St. Louis, MO 63134, USA
 Tel: 877-577-2733
 Fax: 314-427-3324
 Email: service@genomesystems.com

FEATURES

Location/Qualifiers
 1..530
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /cultivar="Columbia Col-0"
 /db_xref="taxon:3702"
 /clone="701676626"
 /tissue_type="inflorescence"
 /dev stage="4 - 7 weeks"
 /clone lib="A. thaliana, Columbia Col-0, inflorescence-1"
 /note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA library was derived from untreated inflorescence tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunt-ended, ligated to SalI adaptors digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

BASE COUNT

163 a 108 c 100 g 158 t 1 others

ORIGIN

Query Match 51.5%; Score 521; DB 9; Length 530;
 Best Local Similarity 99.8%; Pred. No. 1.6e-109;
 Matches 521; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

404 GGGAGAAACATGGGACTTGTCTTCTCTGTTTTCATGATGAGTATTAATTCCTTA 463

Db

522 GGGAGAAACATGGGACTTGTCTTCTCTGTTTTCATGATGAGTATTAATTCCTTA 463

QY

464 CCACACTTAATCTCTACTTGAAGCAATAATGTCAGGATGTCCTTATCAAGCTGCTATG 523

Db

462 CCACACTTAATCTCTACTTGAAGCAATAATGTCAGGATGTCCTTATCAAGCTGCTATG 403

QY

524 TTGCTTCCAAACAGTGAAGATCTCTAGGAGGTATCGTAACAGCCATTCAGATGCAAT 583

Db

402 TTGCTTCCAAACAGTGAAGATCTCTAGGAGGTATCGTAACAGCCATTCAGATGCAAT 343

QY

584 TTCATATCACCCCTCGAAGTGGTTTGCAAAAGAGATGCAATCGATGAATACGATATGCT 643

Db

342 TTCATATCACCCCTCGAAGTGGTTTGCAAAAGAGATGCAATCGATGAATACGATATGCT 283

QY

644 TCTATAAAGATTTAAGCCCGAGGACTGTGTGGTTCACAAGATTGACATCTAGAAAGT 703

Db

282 TCTATAAAGATTTAAGCCCGAGGACTGTGTGGTTCACAAGATTGACATCTAGAAAGT 223

```
QY 704 CATGCCCAAGTACGTAAAGTTTCCGGGAATACACGCCATTAGATGGTGAAGCTATGGTTC 763
Db 222 CATGCCCAAGTACGTAAAGTTTCCGGGAATACACGCCATTAGATGGTGAAGCTATGGTTC 163
QY 764 TGAAGATGCCAACAGAAAGAGAGCTTTTGAATCGGAAAGATCGGAGCTTTGTTAATCT 823
Db 162 TGAAGATGCCAACAGAAAGAGAGCTTTTGAATCGGAAAGATCGGAGCTTTGTTAATCT 103
QY 824 TCTGAGAGACAATACATATGCTCTGATGTTGTAACTTTTACTACCAAAACCTATAAG 883
Db 102 TCTGAGAGACAATACATATGCTCTGATGTTGTAACTTTTACTACCAAAACCTATAAG 43
QY 884 ATTGGCTATTTCGTTCTTATGGATATGATCATCATCTAGT 925
Db 42 ATTGGCTATTTCGTTCTTATGGATATGATCATCATCTAGT 1

RESULT 3
AV527174 520 bp mRNA linear EST 01-SEP-2000
LOCUS AV527174 Arabidopsis thaliana aboveground organs two to six-week
DEFINITION old Arabidopsis thaliana cDNA clone AP230G03R 5', mRNA sequence.
ACCESSION AV527174
VERSION AV527174
KEYWORDS EST.
SOURCE AV527174.1 GI:8686702
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 520)
Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
20363093
JOURNAL 10907847
MEDLINE
PUBMED
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizuekazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1..520
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/cclone="AP230G03R"
/tissue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana aboveground organs two to
six-week old"
/notes="vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 118 a 114 c 121 g 167 t
ORIGIN

Query Match 51.4%; Score 520; DB 9; Length 520;
Best Local Similarity 100.0%; Pred. No. 2.6e-109;
Matches 520; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ATTTAAAGTCATGGCGTCACGTTTATGTTCTCTCTCTCTGTTGCGTGATCCGCGAGC 65
Db 1 ATTTAAAGTCATGGCGTCACGTTTATGTTCTCTCTCTCTGTTGCGTGATCCGCGAGC 60
QY 66 ATTTGCGGAGACGTCATCGAATCATCGATCTCAGAGGAGTTCGATTATTCGCTCT 125
Db 61 ATTTGCGGAGACGTCATCGAATCATCGATCTCAGAGGAGTTCGATTATTCGCTCT 120
QY 126 ATCTCTTCAAGGCGTGAACCTATTGCGGGAACCTCGCATGTTGCTCCAAACACGC 185
```

```
Db 121 ATCTCTTCAATGGCCTGGAAACCTATTGCGCGTGAACCTCGCCATTGTTCTCCAAAACGC 180
QY 186 TTGCTGCAGAGGCTCCGATCTCCAACTCAATTCACAATTCATGGGTATGGCTGACTA 245
Db 181 TTGCTGCAGAGGCTCCGATCTCCAACTCAATTCACAATTCATGGGTATGGCTGACTA 240
QY 246 TAACGATGGTTCGTGGCTTCATGTTGTTATCCATCTGACTTTTAAAGAGAGAGATTC 305
Db 241 TAACGATGGTTCGTGGCTTCATGTTGTTATCCATCTGACTTTTAAAGAGAGAGATTC 300
QY 306 AACCTTGATGGATGGTCTTGAGAACTAGTGGCTAGTCTCAGTTGTGGTTCTCCATCATC 365
Db 301 AACCTTGATGGATGGTCTTGAGAACTAGTGGCTAGTCTCAGTTGTGGTTCTCCATCATC 360
QY 366 ATGCAATGGTGGAAAGGCTCATTTTGGGCCAGGATGGGAGAAACATGGACTTCTTC 425
Db 361 ATGCAATGGTGGAAAGGCTCATTTTGGGCCAGGATGGGAGAAACATGGACTTCTTC 420
QY 426 TTCTCTCTGTTTTCATGATGAGTATATTAATCTTCTTACCACACTTAATCTCTACTTGA 485
Db 421 TTCTCTCTGTTTTCATGATGAGTATATTAATCTTCTTACCACACTTAATCTCTACTTGA 480
QY 486 GCATAATGTCACGATGTCCTTTATCAAGCTGGCTATGTT 525
Db 481 GCATAATGTCACGATGTCCTTTATCAAGCTGGCTATGTT 520

RESULT 4
AV832175 607 bp mRNA linear EST 01-APR-2002
LOCUS AV832175 RAF111 Arabidopsis thaliana cDNA clone RAF111-05-B04 5',
DEFINITION mRNA sequence.
ACCESSION AV832175
VERSION AV832175.1 GI:19874235
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 607)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Cono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shingawa,A., Muramatsu,M., Hayashizaki,Y.
and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekic@tc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified lambdaBamBla-F10-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
details.
Location/Qualifiers
1..607
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF111-05-B04"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_hosts="DH10B"
/clone_lib="RAF111"
/notes="Site 1: BamHI; Site 2: SalI; subjected to various
treatments (dehydration, cold, high salt, ABA, heat and UV
```

```

Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.

FEATURES             source
Location/Qualifiers
1. 490
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="AP230g03_f"
/tissue_type="aboveground organ"
/dev_stage="two to six-week old plants"
/clone_lib="Arabidopsis thaliana Above-ground organ from
two to six-week old plants Columbia"
/note="vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT      147 a      99 c      87 g      157 t
ORIGIN
Query Match      48.4%; Score 490; DB 9; Length 490;
Best Local Similarity 100.0%; Pred. No. 2.2e-102;
Matches 490; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      485  AGCATATGTCACGGATGTCCTTTATCAAGCTGGCTATGTTGCTTCCAAACAGTGAAGAAGT 544
Db      490  AGCATATGTCACGGATGTCCTTTATCAAGCTGGCTATGTTGCTTCCAAACAGTGAAGAAGT 431
QY      545  ATCCCTCTAGGAGGTATCGTAAACAGCCATTGAGATGATGATTCATATCACCCCTGAAGTGG 604
Db      430  ATCCCTCTAGGAGGTATCGTAAACAGCCATTGAGATGATGATTCATATCACCCCTGAAGTGG 371
QY      605  TTTCGAAAGAGATGCAATCGATGAATACGTATATGCTTCTATAAAGATTTTAAGCCCA 664
Db      370  TTTCGAAAGAGATGCAATCGATGAATACGTATATGCTTCTATAAAGATTTTAAGCCCA 311
QY      665  GGGAGCTGTTGGTTTCAACAGATTGACATCTAGAAGTCTATGCTGGAAGTGCACCAAGTGAAGT 724
Db      310  GGGAGCTGTTGGTTTCAACAGATTGACATCTAGAAGTCTATGCTGGAAGTGCACCAAGTGAAGT 251
QY      725  TGCCGGATACACGCCATTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 784
Db      250  TGCCGGATACACGCCATTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 191
QY      785  AAGCTCTTTGAATCGGAAAGATGGAGCTTTGTTATCTTCTGAGAGACATACATACAT 844
Db      190  AAGCTCTTTGAATCGGAAAGATGGAGCTTTGTTATCTTCTGAGAGACATACATACAT 131
QY      845  GTCTCTGATGTTGTAACCTTACTACCAAACTATTAAGATGCTTATTTGCTTCTATT 904
Db      130  GTCTCTGATGTTGTAACCTTACTACCAAACTATTAAGATGCTTATTTGCTTCTATT 71
QY      905  GGATATGATCATCTATTACTGTTAAATCAAGTTTCTTTCTAATAATGTAAGATCAGAA 964
Db      70  GGATATGATCATCTATTACTGTTAAATCAAGTTTCTTTCTAATAATGTAAGATCAGAA 11
QY      965  AATCCATAAG 974
Db      10  AATCCATAAG 1

RESULT 6
AB038725/c
LOCUS
DEFINITION
AB038725 Arabidopsis thaliana aboveground organs two to six-week
old Arabidopsis thaliana cDNA clone AP232g06f 3', mRNA sequence.
ACCESSION
AB038725
VERSION
AB038725.1
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
JOURNAL
MEDLINE
20363093
PUBMED
10907847
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research

```

```

). Dark-grown plants"
BASE COUNT      148 a      134 c      134 g      190 t      1 others
ORIGIN
Query Match      48.5%; Score 491; DB 9; Length 607;
Best Local Similarity 93.5%; Pred. No. 1.3e-102;
Matches 548; Conservative 0; Mismatches 1; Indels 37; Gaps 2;

QY      1  ATCGAATTAAAGTCAATGGC-GTCAGGTTATGCTTCTCCTTCTCGTTCGCTGATCGC 59
Db      22  ATCGAATTAAAGTCAATGGCAGTCAGCTTATGCTTCTCCTTCTCGTTCGCTGATCGC 81
QY      60  CGGAGCATTTGCCGAGAGCTGATCGAATCAATCGATCTCAGAGGAGTTCGATTATT 119
Db      82  CGGAGCATTTGCCGAGAGCTGATCGAATCAATCGATCTCAGAGGAGTTCGATTATT 141
QY      120  CGCTCTATCTCTTCAATGGCTCGAACTTATGCCGCGGAATCGCCATGTTGCTCCAA 179
Db      142  CGCTCTATCTCTTCAATGGCTCGAACTTATGCCGCGGAATCGCCATGTTGCTCCAA 201
QY      180  AAACGCTTGTGAGAGGCTCCGATGCTCAACTCAATTCACAAATTCATGGGTATGGCC 239
Db      202  AAACGCTTGTGAGAGGCTCCGATGCTCCAACTCAATTCACAAATTCATGGGTATGGCC 261
QY      240  TGACTATAACGATGTTTGGTGGCTTCATGTTGTTATCGATCTGACTTTAAAGAGAAGA 299
Db      262  TGACTATAACGATGTTTGGTGGCTTCATGTTGTTATCGATCTGACTTTAAAGAGAAGA 321
QY      300  GATTTCAGGTTGATGATGCTTGAAGTACTGCGCTAGCTCAGTCTGTTGTTCTCC 359
Db      322  GATTTCAGGTTGATGATGCTTGAAGTACTGCGCTAGCTCAGTCTGTTGTTCTCC 381
QY      360  ATCATCATGCAATGGTGGGAAAGGTCATTTTGGGGCCAGAGTGGG----- 406
Db      382  ATCATCATGCAATGGTGGGAAAGGTCATTTTGGGGCCAGAGTGGGAGTTCATGCTTA 441
QY      407  -----AGAAACATGGGACTGTTCTTCTCCCTGTTTTCATGA 443
Db      442  ATCTCACAACTTTTGCAATACAGAAACATGGGACTGTTCTTCTCCCTGTTTTCATGA 501
QY      444  TGAGTATAATTAATCTTCTTACCACATTAATCTTCTTGAAGCATAATCTCAGGATG 503
Db      502  TGAGTATAATTAATCTTCTTACCACATTAATCTTCTTGAAGCATAATCTCAGGATG 561
QY      504  CCTTTATCAAGCTGGCTATGTTGCTTCCAAACAGTGAAGAATTCCT 549
Db      562  CCTTTATCAAGCTGGCTATGTTGCTTCCAAACAGTGAAGAATTCCT 607

RESULT 5
AB038725/c
LOCUS
DEFINITION
AB038725 Arabidopsis thaliana Above-ground organ from two to
six-week old plants Columbia Arabidopsis thaliana cDNA clone
AP230g03_f 3', mRNA sequence.
ACCESSION
AB038725
VERSION
AB038725.1
KEYWORDS
Arabidopsis thaliana (thale cress)
SOURCE
Arabidopsis thaliana
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7, 175-180 (2000)
JOURNAL
MEDLINE
20363093
PUBMED
10907847
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research

```


934 AGTTCTTTCTTAATATGTAAGATCAGAAATCCATAAGAGATATCAACATTTGAGT 993
 62 AGTTTCTTTCTTAATATGTAAGATCAGAAATCCATAAGAGATATCAACATTTGAGT 3
 994 TC 995
 2 CC 1

RESULT 8
 AV820692/c
 LOCUS
 DEFINITION AV820692 419 bp mRNA linear EST 01-APR-2002
 AV820692 RAF111 Arabidopsis thaliana cDNA clone RAF111-11-M08 3',
 mRNA sequence.
 ACCESSION AV820692 GI:19862671
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 419)
 REFERENCE Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
 and Shinozaki, K.
 TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
 JOURNAL Unpublished
 COMMENT Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: msekierc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified pBluescript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
 details.

FEATURES
 source
 1..419
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF111-11-M08"
 /dev_stage="plants at various developmental stages from
 germination to mature seeds"
 /lab_host="DH10B"
 /clone_lib="RAF111"
 /note="Site 1: BamHI; Site 2: SalI; subjected to various
 treatments (dehydration, cold, high salt, ABA, heat and UV
 treatments)"
). Dark-grown plants"
 BASE COUNT 128 a 85 c 134 t
 ORIGIN

Query Match 41.2%; Score 417.4; DB 9; Length 419;
 Best Local Similarity 99.8%; Pred. No. 1.2e-85;
 Matches 418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

542 AGTATCTCTAGGAGGTATCGTACGCCATTCAGATGATTCATATCACCCTCGAAG 601
 419 AGTATCTCTAGGAGGTATCGTACGCCATTCAGATGATTCATATCACCCTCGAAG 360

602 TGGTTTGCAAAAGAGATCGATGAAATACGTATATGCTTCTATAAAGATTTTAAAGC 661
 359 TGGTTTGCAAAAGAGATCGATGAAATACGTATATGCTTCTATAAAGATTTTAAAGC 300

662 CCAGGGACTGTGTTGGTTCCAAAGATTGACATCTAGAAAGTCATGCCCAAGTACGTAA 721

299 CCAGGGACTGTGTTGGTTCCAAAGATTGAAATCTAGAAAGTCATGCCCAAGTACGTAA 240
 722 GTTTGCGGGAATACAGCCATTAGATGGTGAAGCTATGTTCTGAAGATGCCCAACAGAAA 781
 239 GTTTGCGGGAATACAGCCATTAGATGGTGAAGCTATGTTCTGAAGATGCCCAACAGAAA 180
 782 GAGAAGCTCTTTGAATCGGAAAAGATGGGAGCTTTGTTATCTCTCTGAGAGACAATACATA 841
 179 GAGAAGCTCTTTGAATCGGAAAAGATGGGAGCTTTGTTATCTCTCTGAGAGACAATACATA 120
 842 CATGCTCTGATGTTGTAACCTTTTACTACCAAAACCTATAAAGATTTGCTTATTTGCTTCT 901
 119 CATGCTCTGATGTTGTAACCTTTTACTACCAAAACCTATAAAGATTTGCTTATTTGCTTCT 60
 902 ATTGGATATGATCATCATCTACTGTTAAATCAAGTTCTTTCTTAATAATGTAGAAGATC 960
 59 ATTGGATATGATCATCATCTACTGTTAAATCAAGTTCTTTCTTAATAATGTAGAAGATC 1

AV819666 418 bp mRNA linear EST 01-APR-2002
 AV819666/c
 LOCUS
 DEFINITION AV819666 RAF111 Arabidopsis thaliana cDNA clone RAF111-06-E06 3',
 mRNA sequence.
 ACCESSION AV819666 GI:19861609
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 418)
 REFERENCE Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
 and Shinozaki, K.
 TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
 JOURNAL Unpublished
 COMMENT Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: msekierc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified pBluescript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further
 details.

FEATURES
 source
 1..418
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF111-06-E06"
 /dev_stage="plants at various developmental stages from
 germination to mature seeds"
 /lab_host="DH10B"
 /clone_lib="RAF111"
 /note="Site 1: BamHI; Site 2: SalI; subjected to various
 treatments (dehydration, cold, high salt, ABA, heat and UV
 treatments)"
). Dark-grown plants"
 BASE COUNT 128 a 85 c 74 g 131 t
 ORIGIN

Query Match 41.1%; Score 416.4; DB 9; Length 418;
 Best Local Similarity 99.8%; Pred. No. 2e-85;
 Matches 417; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

BASE COUNT	117 a	80 c	63 g	133 t	1 others
ORIGIN					
Query Match	38.7%; Score 391.4; DB 14; Length 394;				
Best Local Similarity	99.5%; Pred. No. 1.1e-79;				
Matches 392; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
QY	599	AAAGTGGTTGCAAAAGAGATGCAATCGATGAAATACGTATATGCTTCTTATAAAGATTTTA	658		
Db	394	AAAGNGTTTGCAAAAGAGATGCAATCGATGAAATACGTATATGCTTCTTATAAAGATTTTA	335		
QY	659	AGCCAGGAGCTGTGTGTTTCAAGATTTGACATCTAGAAAGTCAATGCCCCAAGTACG	718		
Db	334	AGCCAGGAGCTGTGTGTTTCAAGATTTGACATCTAGAAAGTCAATGCCCCAAGTACG	275		
QY	719	TAAGTTTCCGGAATACACGCAATPAGATGTGAAGTATGTTCTGAAAGATGCCAAG	778		
Db	274	TAAGTTTCCGGAATACACGCAATPAGATGTGAAGTATGTTCTGAAAGATGCCAAG	215		
QY	779	AAAGAGAGCTCTTTGAATCGGAAAGATGGAGCTTTGTTATCTCTGAGAGACAATAC	838		
Db	214	AAAGAGAGCTCTTTGAATCGGAAAGATGGAGCTTTGTTATCTCTGAGAGACAATAC	155		
QY	839	ATACATGTCTGTGATGTTGTAATCTTACTACAAAGCTATAAAGATGGCTTATTTGCT	998		
Db	154	ATACATGTCTGTGATGTTGTAATCTTACTACAAAGCTATAAAGATGGCTTATTTGCT	95		
QY	899	TCATTGGATATGATATCAATCTACTGGTAATCAAGTTTCTTCTTAATATGTAAGA	958		
Db	94	TCATTGGATATGATATCAATCTACTGGTAATCAAGTTTCTTCTTAATATGTAAGA	35		
QY	959	TCAGAAATCCATAAGAGATATCAACATTTGAG	992		
Db	34	TCAGAAATCCATAAGAGATATCAACATTTGAG	1		
RESULT 12					
LOCUS	BQ791108				
DEFINITION	E3954 Chinese cabbage etiolated seedling library Brassica rapa subsp. pekinensis cDNA clone E3954, mRNA sequence.				
ACCESSION	BQ791108				
VERSION	BQ791108.1				
KEYWORDS	EST.				
SOURCE	Brassica rapa subsp. pekinensis				
ORGANISM	Brassica rapa subsp. pekinensis				
REFERENCE	Ryu,S.H., Yang,K.A., Lee,S.Y., Kim,H.-I., Cho,M.J. and Lim,C.O. Expressed Sequence Tags of Chinese Cabbage Etiolated Seedling cDNA (2002)				
AUTHORS	Unpublished				
TITLE	Contact: Lim, C.O.				
JOURNAL	Plant Molecular Biology & Biotechnology Research Centre				
COMMENT	Gyeongsang National University #900 Gazwa-dong, Jinju 660-701, Korea				
FEATURES	Tel: 82 55 751 6255				
source	Fax: 82 55 759 9363				
	Email: colimanongae.gsmu.ac.kr				
	Seq primer: 17.				
	Location/Qualifiers				
	1..570				
	/organism="Brassica rapa subsp. pekinensis"				
	/mol_type="mRNA"				
	/cultivar="Jangwon"				
	/db_xref="taxon:51351"				
	/clone="E3954"				
	/tissue_type="Etiolated seedling"				
	/lab_host="XL-1 Blue"				
	/clone_lib="Chinese cabbage etiolated seedling library"				
	/note="Vector: pSPORT 1; Site_1: Sal I; Site_2: Not I"				

BASE COUNT	122 a	149 c	133 g	166 t
ORIGIN				
Query Match	37.4%; Score 378.8; DB 13; Length 570;			
Best Local Similarity	89.6%; Pred. No. 9.5e-77;			
Matches 407; Conservative 0; Mismatches 47; Indels 0; Gaps 0;				
QY	76	GACGTCACTCGAAGCTCAATCGATCTCAGAGGAGTTCGATTATTTTCGCTCTATCTCTCAA	135	
Db	117	GACGTGTCGCTACCAATCGCTCGCAGAGGAGTTTGATTATTTTCGCTCTCTCTCCAA	176	
QY	136	TGGCTCGAAGCTATTGCGGTGGAACTTCGCCATTGTTGCTCCAAAAACGCTTCTCTCGAGA	195	
Db	177	TGGCCCGGAACCTACTCCGCTGGAACTTCGCCATTGCTGCTCCAAAAACGCTTCTCTAGA	236	
QY	196	GGCTCCGATGCTCCAACTCAATTCACAATTCATGGGTATGGCTGACTATACGATGGT	255	
Db	237	GGCTCGGACACTCCAACTCAATTCACCATCCATGGACTATGGCGGAGCTATAACGACGGT	296	
QY	256	TCGTGGCTCTTCATGTTGTTATCGATCTGACTTTTAAAGAGAAGGAGATTTTCAACGTTGATG	315	
Db	297	TCCTGGCTCTTCATGTTGTTATCGATCCGACTTCAATGAGAAGGAGATCTCTACGTTGATG	356	
QY	316	GATGGTCTTGAGAGTACTGGCTAGTCTCAGTTGTTGTTCTCCATCANCATCGAATGGT	375	
Db	357	GATGGCATGAGTAAGTACTGGCCAGTCTCAGCTGGTTTCTCCGTCATCTTCCCATGGT	416	
QY	376	GGGAAAGGGTCATTTTGGGGCCACGAGTGGGGAAGAACATGGGACTTGTCTTCTCTGTT	435	
Db	417	GGAAAGGATCATTTTGGGGCCATGAGTGGGGAAGAACATGGGACTTGTCTTCTCTGTT	476	
QY	436	TTTTCATCATGATGATTAATTTACTTCTTACCACACTTAATCTCTACTTTGAGCATATGTC	495	
Db	477	ATTCTGATGAGTATAGTTACTTCTTACCACGCTTAATCTATCTTAAAGCATATGTC	536	
QY	496	ACGGATGTCCTTTATCAAGCTGGCTATGTTGCTT	529	
Db	537	ACGGATGTCCTTTATCAAGCTGGCTATGTTGCTT	570	
RESULT 13				
LOCUS	AV816618			
DEFINITION	AV816618 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-91-F09 3', mRNA linear EST 01-APR-2002			
ACCESSION	AV816618			
VERSION	AV816618.1			
KEYWORDS	EST.			
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids ; eurosids II; Brassicales; Brassicaceae; Arabidopsi			
AUTHORS	Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.			
TITLE	Large scale analysis of Arabidopsis full-length cDNA (2002b)			
JOURNAL	Unpublished			
COMMENT	Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: msek@rtc.riken.go.jp			
FEATURES	An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further			

BASE COUNT	122 a	149 c	133 g	166 t
ORIGIN				
Query Match	37.4%; Score 378.8; DB 13; Length 570;			
Best Local Similarity	89.6%; Pred. No. 9.5e-77;			
Matches 407; Conservative 0; Mismatches 47; Indels 0; Gaps 0;				
QY	76	GACGTATCGAATCAATCGATCTCAGAGGAGTTCGATTTATTTCTCTCTCTTCAA	135	
DB	117	GACGTCTGCTACTCAATCGCTCGCAGAGGAGTTTGATTTCTCTCTCTCTCAA	176	
QY	136	TGGCTCGAACCATTATGCGTGAACCTCGCCATTGTTGCTCCAAAACGCTTCTGCAGA	195	
DB	177	TGGCCCGAACCATTATGCGTGAACCTCGCCATTGTTGCTCCAAAACGCTTCTGTAGA	236	
QY	196	GGCTCCGATCTCCAACTCAATTCACAATTCATGGTTATGGCTGACTATTAACGATGT	255	
DB	237	GGCTCCGACACTCCAACTCAATTCACAATTCATGGTTATGGCTGACTATTAACGATGT	296	
QY	256	TCGTGGCTCTCATGTTGTTATTCGATCTGACTTTTAAAGAGAGAGATTTTCAACGTTGATG	315	
DB	297	TCGTGGCTCTCATGTTGTTATTCGATCTGACTTTTAAAGAGAGAGATTTTCAACGTTGATG	356	
QY	316	GATGGTCTTGAGAGTACTGGCTGCTAGTCTCAGTTGTTCTCCATCATCATGCAATGT	375	
DB	357	GATGGCATGATAAGTACTGGCCAGTCTCAGCTGTGGTTCTCCGTCATCTTGGCATGT	416	
QY	376	GGGAAGGGCTCATTTTGGGCCACGAGTGGGAGACATGGGACTTGTCTTCTCTCTGT	435	
DB	417	GGAAAGAGTCAATTTTGGGCCCATGAGTGGGAGAAACATGGGACTTGTCTTCTCTGT	476	
QY	436	TTTCATGATGATATAATTACTTCTTACCACATTTAATCTCTTCTTGAAGCATATGTC	495	
DB	477	ATTCGTGATGATATAATTACTTCTTACCACATTTAATCTATCTATTAAGCATATGTC	536	
QY	496	ACGGATGCTCTTTATCAAGCTGGCTATGTTGCTT	529	
DB	537	ACGGATGCTCTTTATCAAGCTGGCTATGTTGCTT	570	
RESULT 13				
LOCUS	AV816618			
DEFINITION	AV816618 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-91-P09 3', mRNA sequence.			
ACCESSION	AV816618			
VERSION	AV816618.1			
KEYWORDS	EST.			
SOURCE	Arabidopsis thaliana (thale cress)			
ORGANISM	Arabidopsis thaliana			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids			
AUTHORS	; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
TITLE	1 (bases 1 to 400)			
JOURNAL	Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Ono,K., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.			
COMMENT	Large scale analysis of Arabidopsis full-length cDNA (2002b)			
	Unpublished			
	Contact: Motoaki Seki			
	Plant Functional Genomics Research Group			
	RIKEN Genomic Sciences Center			
	3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan			
	Tel: 81-298-36-4359			
	Fax: 81-298-36-9060			
	Email: meeki@rtr.riken.go.jp			
	An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pluscript vector. Please visit our web site (http://www.gen.riken.go.jp/e/plant/index_e.html) for further			

details.		Location/Qualifiers	
FEATURES	source	1..400	
		/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/db_xref="taxon:3702"	
		/clone="RAF109-91-F09"	
		/dev stage="plants at various developmental stages from germination to mature seeds"	
		/lab_host="DH10B"	
		/lab_host="DH10B"	
		/clone_lib="RAF109"	
		/notes="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) treatments"	
BASE COUNT		127 a	64 g 132 t
ORIGIN		77 c	
Query Match			
Best Local Similarity 37.1%; Score 375.6; DB 9; Length 400;			
Matches 378; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
QY	628	GAATACGATATGCTTCATTAAGATTTTAAGCCAGGAGCTGTGTTGTTTCAACAAGAT	687
Db	400	GAATACGATATGCTTCATTAAGATTTTAAGCCAGGAGCTGTGTTGTTTCAACAAGAT	341
QY	688	TTGACATCTAGAAGTCATGCCCAAGTAGCTTAAGTTTCCCGAATACACGCCATTAGAT	747
Db	340	TTGACATCTAGAAGTCATGCCCAAGTAGCTTAAGTTTCCCGAATACACGCCATTAGAT	281
QY	748	GGTGAAGCTATGCTTCGAGAGTCCCAAGAGAGAGAGCTCTTTGAATCGAAAGAT	807
Db	280	GGTGAAGCTATGCTTCGAGAGTCCCAAGAGAGAGAGCTCTTTGAATCGAAAGAT	221
QY	808	GGGAGCTTTGTTATCTTCTGAGAGACAATACATACATGCTCTGATGTTGTAACCTTACT	867
Db	220	GGGAGCTTTGTTATCTTCTGAGAGACAATACATACATGCTCTGATGTTGTAACCTTACT	161
QY	868	ACCAAACTTAAAGATTGGCTTATTTGTTCTATGATGATGATCATCTACTGCT	927
Db	160	ACCAAACTTAAAGATTGGCTTATTTGTTCTATGATGATGATCATCTACTGCT	101
QY	928	AAATCAAGTTCTTTCTTAATATGTAAGATCAGAAATCCATAAGAGATATCAACAT	987
Db	100	AAATCAAGTTCTTTCTTAATATGTAAGATCAGAAATCCATAAGAGATATCAACAT	41
QY	988	TTGAGTCTATGTTAAATAAA 1009	
Db	40	TTGAGTCTATGTTAAATTGAA 19	
RESULT 14			
AV527256 371 bp mRNA linear EST 01-SEP-2000			
LOCUS AV527256 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone AP232906 5', mRNA sequence.			
DEFINITION AV527256			
ACCESSION AV527256.1 GI:8686784			
VERSION			
KEYWORDS			
SOURCE Arabidopsis thaliana (thale cress)			
ORGANISM Arabidopsis thaliana			
REFERENCE			
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.			
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries			
JOURNAL DNA Res. 7, 175-180 (2000)			
MEDLINE 20363093			
PUBMED 10907847			
COMMENT Contact: Erika Asamizu			
The First Laboratory for Plant Gene Research			
FEATURES		Location/Qualifiers	
source	1..371		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/strain="Columbia"		
	/db_xref="taxon:3702"		
	/clone="AP232906R"		
	/tissue_type="aboveground organs"		
	/dev stage="two to six-week old"		
	/clone_lib="Arabidopsis thaliana aboveground organs two to six-week old"		
	/note="vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"		
BASE COUNT		82 a	85 c 90 g 114 t
ORIGIN			
Query Match			
Best Local Similarity 36.7%; Score 371; DB 9; Length 371;			
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	10	AAGTCAATGGCGTCACCGTTTATGTCTTCTCTCTCGTGTATCGCGGAGCATTT	69
Db	1	AAGTCAATGGCGTCACCGTTTATGTCTTCTCTCTCGTGTATCGCGGAGCATTT	60
QY	70	GCCGAGAGCTCATCGAATCAATCGATCTCAGAGGAGTTCGATTATTTTCGCTCTATCT	129
Db	61	GCCGAGAGCTCATCGAATCAATCGATCTCAGAGGAGTTCGATTATTTTCGCTCTATCT	120
QY	130	CTTCAATGGCGTGGAACTTATGGCGTGAACCTGCGCATTTGTTCTCCAAAAACGTTGC	189
Db	121	CTTCAATGGCGTGGAACTTATGGCGTGAACCTGCGCATTTGTTCTCCAAAAACGTTGC	180
QY	190	TGCAGAGCTCCGAGTCCAACTCAATTCATCAATTCATGGTTATGGCTGATATAAC	249
Db	181	TGCAGAGCTCCGAGTCCAACTCAATTCATCAATTCATGGTTATGGCTGATATAAC	240
QY	250	GATGTTGCGTGGCTTCATGTTTATCGATCTCAGCTTTTAAAGAGAGGAGATTCAACG	309
Db	241	GATGTTGCGTGGCTTCATGTTTATCGATCTCAGCTTTTAAAGAGAGGAGATTCAACG	300
QY	310	TTGATGATGCTTGTGAGAGTACTGCGCTAGTCTCAGTTGTGTTTCCATCATATGC	369
Db	301	TTGATGATGCTTGTGAGAGTACTGCGCTAGTCTCAGTTGTGTTTCCATCATATGC	360
QY	370	AATGTTGGGAA 380	
Db	361	AATGTTGGGAA 371	
RESULT 15			
R90379 453 bp mRNA linear EST 30-DEC-1997			
LOCUS R90379 16734 Lambda-PRL2 Arabidopsis thaliana cDNA clone 18501877, mRNA sequence.			
DEFINITION R90379			
ACCESSION R90379.1 GI:957919			
VERSION			
KEYWORDS			
SOURCE Arabidopsis thaliana (thale cress)			
ORGANISM Arabidopsis thaliana			
REFERENCE			
AUTHORS Newman, T., deBruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C.			
TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones			
JOURNAL Plant Physiol. 106, 1241-1255 (1994)			
MEDLINE 95148729			

266 CATGTGTTATCGATGCTGCTTTAAAGAGAGAGGAGATTTCACCTTGATGATGCTGTTG 325
281 CATGTGTTCCAGGTGCTGCTTATGATGAACAACAAGATTTCACCATTTGCTTGATGCTTTGC 340
326 AGAAGTACTGCGCTAGTCTCACTGCTGCTGCTTCCATCATCATGCAATGGTGGGAAGGGT 385
341 ACACATATATGCGGATACCTAAGTCTGAACGGTATAGAAATCATGGATACAAAGCAAGT 400
386 CAATTTGGGGCCACGAGTGGGAGAACATGGGACTTGTCTTCTCCCTGTTTTCATGATG 445
401 CGTTTGGGCTTATCAATGGGAGCAACATGATGATGTCGACCTTAGTAACAGGGGACC 460
446 AGTATAATTAATCTCTTACCACACTTAATCTCTACTTGAACATATAATGTCACGATGTC 505
461 AATATGGATATTTCTGACAGCCTTGAATTTGATTTGGAGTATAAATGTCAGCGAGTTC 520
506 TTTATCAAGCTGCTATGTTCTCCACAGTGAAGATATCTCTAGGAGGTATGTTAA 565
521 TATCAAAAGCTGGAATTAACCTTCAAAATCAAGACTTACTCATCTGCTGCCATTTGTT 580
566 CAGCATTCAGAAATCATTTTCATATCACCTTGAAGTGGTTTGCAAAA---GAGATGCAA 622
581 CTACTATTGAAGTCTTCTGGTGCAACCGCAATTTGGTCTGCAAAAATGGTGTATGTA 640
623 TCGATGAATAGTATGCTTCTATAGATTTTAAAGCTTAAAGCCAGGAGTGTGTTG 676
641 TTAAGAAGTCCGCTTGTTGTTTACCAGACCAAGATCCGAGAGTGTGTGG 694

RESULT 2

US-08-986-963-1
; Sequence 1, Application US/08986963
; Patent No. 5958730
; GENERAL INFORMATION:
; APPLICANT: Rosteck Jr., Paul R.
; TITLE OF INVENTION: Streptococcus Pneumoniae Gene Sequence
; TITLE OF INVENTION: Ftsy
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: US
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/986,963
; FILING DATE: December 8, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317/276-3334
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1278 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1278

US-08-986-963-1

Query Match 3.9%; Score 39.4; DB 2; Length 1278;
Best Local Similarity 52.1%; Pred. No. 0.13;
Matches 88; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 513 AGCTGGCTATGTTGCTTCCACAGTGAAGATATCTCTAGGAGGTATCGTAACAGCCAT 572
Db 1104 AACGGAAATGTTTGTACTAAGATTGATGGAATGCTCGAGGAGGTGTGTTCTAGCCAT 1163
QY 573 TCAGAATGCATTTTCATATCACCCCTGAAAGTGGTTTGCACAAAGAGATGCAATCGATGAAT 632
Db 1164 TCGTGAAGAACTCAATATCTCTGTAATAATTTGATTTGGTGAATAAATCGATCATAT 1223
QY 633 ACGTATATGCTTCTATAAGATTTTAAAGCCAGGAGTGTGTTGTTCA 681
Db 1224 TGGAGAGTTTAACTCAGAAATCTTTATGAAGTCTCTTTGGAGGGTTTA 1272

RESULT 3

US-08-986-963-3
; Sequence 3, Application US/08986963
; Patent No. 5958730
; GENERAL INFORMATION:
; APPLICANT: Rosteck Jr., Paul R.
; TITLE OF INVENTION: Streptococcus Pneumoniae Gene Sequence
; TITLE OF INVENTION: Ftsy
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: US
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/986,963
; FILING DATE: December 8, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11763
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317/276-3334
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1278 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-986-963-3

Query Match 3.9%; Score 39.4; DB 2; Length 1278;
Best Local Similarity 35.5%; Pred. No. 0.13;
Matches 60; Conservative 28; Mismatches 81; Indels 0; Gaps 0;
QY 513 AGCTGGCTATGTTGCTTCCACAGTGAAGATATCTCTAGGAGGTATCGTAACAGCCAT 572
Db 1104 AACGGAAATGTTTGTACTAAGATTGATGGAATGCTCGAGGAGGTGTGTTCTAGCCAU 1163
QY 573 TCAGAATGCATTTTCATATCACCCCTGAAAGTGGTTTGCACAAAGAGATGCAATCGATGAAT 632
Db 1164 UCGGAAGAACUAAUUAUCCUGAUAUUGUUGUGAUAUUGUGUGAUAUUAUUGAUAU 1223

QY 633 ACCTATATGCTTCTATATAAGATTAAAGCCAGGAGCTGTGTGTTCA 681
 Db 1224 UGAGAGUUUUAACUCAGAAAACUUUAUGAAAGGUCUCUUGGAAGGUUA 1272

RESULT 4

US-08-986-963-4

; Sequence 4, Application US/08986963

; Patent No. 5958730

; GENERAL INFORMATION:

; APPLICANT: Rostek Jr., Paul R.

; TITLE OF INVENTION: Streptococcus Pneumoniae Gene Sequence

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Eli Lilly and Company

; STREET: Lilly Corporate Center

; CITY: Indianapolis

; STATE: Indiana

; COUNTRY: US

; ZIP: 46285

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/986,963

; FILING DATE: December 8, 1997

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Webster, Thomas D.

; REGISTRATION NUMBER: 39,872

; REFERENCE/DOCKET NUMBER: X-11763

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 317/276-3334

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2764 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-986-963-4

Query Match 3.9%; Score 39.4; DB 2; Length 2764;

Best Local Similarity 52.1%; Pred. No. 0.18;

Matches 88; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 513 AGCTGGCTATGTTGCTTCCACAGTGAAAAGTATCCTCTAGGAGGTATCGTAACAGCCAT 572

Db 1450 AACGGGAATTGTTTGACTAAGATTGATGGAACCTCTCGAGAGGTGTGTTCTAGCCAT 1509

QY 573 TCAGATGATTCATATACCCCTGAGTGGTTTGCAGAAAGATGCAATCGATGAAT 632

Db 1510 TCGTGAAGAACCAATATCTCTGTAATTTGATGTTTGGTGAAGAAATCGATGATAT 1569

QY 633 ACCTATATGCTTCTATATAAGATTAAAGCCAGGAGCTGTGTGTTCA 681

Db 1570 TGGAGAGTTTAACTCAGAAAACCTTATGAAGGTCCTCTTGAAGGTTTA 1618

RESULT 5

US-08-961-527-126/c

; Sequence 126, Application US/08961527

; Patent No. 6420135

; GENERAL INFORMATION:

; APPLICANT: Charles Kunach

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.4mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961,527

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 126:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 13121 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

US-08-961-527-126

Query Match 3.9%; Score 39.4; DB 4; Length 13121;

Best Local Similarity 52.1%; Pred. No. 0.34;

Matches 88; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 513 AGCTGGCTATGTTGCTTCCACAGTGAAAAGTATCCTCTAGGAGGTATCGTAACAGCCAT 572

Db 4788 AACGGGAATTGTTTGACTAAGATTGATGGAACCTCTCGAGAGGTGTGTTCTAGCCAT 4729

QY 573 TCAGATGATTCATATACCCCTGAGTGGTTTGCAGAAAGATGCAATCGATGAAT 632

Db 4728 TCGTGAAGAACCAATATCTCTGTAATTTGATGTTTGGTGAAGAAATCGATGATAT 4669

QY 633 ACCTATATGCTTCTATATAAGATTAAAGCCAGGAGCTGTGTGTTCA 681

Db 4668 TGGAGAGTTTAACTCAGAAAACCTTATGAAGGTCCTCTTGAAGGTTTA 4620

RESULT 6

US-08-903-325-3/c

; Sequence 3, Application US/08903325

; Patent No. 6262338

; GENERAL INFORMATION:

; APPLICANT: Peter Schreier; Thomas Herget; and

; APPLICANT: Jeff Schell

; TITLE OF INVENTION: RESISTANCE GENES

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SPRUNG HORN KRAMER & WOODS

; STREET: 660 White Plains Road

; CITY: Tarrytown

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10591-5144

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb

; MEDIUM TYPE: storage

; COMPUTER: NEC Powermate SX/20

; OPERATING SYSTEM: DOS

; SOFTWARE: Wordperfect 5.1

; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/903,325
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/383,747
; FILING DATE: 27-JAN-1995
; APPLICATION NUMBER: DE 40 31 758.7
; FILING DATE: 06-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/766990
; FILING DATE: 27-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/235106
; FILING DATE: 28-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Kurt G. Briscoe
; REGISTRATION NUMBER: 33,141
; REFERENCE/DOCKET NUMBER: Bayer 8297.2-KGB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (914) 332-1700
; TELEFAX: (914) 332-1844
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3412 nucleotides
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Peanut (arachis hypogaea)
; US-08-903-325-3

Query Match 3.8%; Score 38.2; DB 3; Length 3412;
Best Local Similarity 49.3%; Pred. No. 0.42;
Matches 100; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 756 TATGGTTCAGATGCCAACAGAGAGAGCTTTTGAATCGAAAGATGGAGCTT 815
DB 2386 TTTTGTCATGAATTAATAATAATATATAGGATTTTAATTTTGACACTAATATGT 2327

QY 816 TGTATCTCTCAGAGACATACATACATGCTCTGATGTTGTAACTTACTACCAAAAC 875
DB 2326 AATAGATATTGTGATAATTTATATATATTTAGATAATTTTTTAATTTTATATACTAAC 2267

QY 876 CTATAAGATTCGCTTATTTCTGTTCTATTTGGATATGATCATCTACTGTAAATCAAG 935
DB 2265 AATGTAAATTAATTAATTTCTATTGATATAACATACACAAATTAATATGTGTAAAAA 2207

QY 936 TTTCTTTCTAATAATGTAGAAGA 958
DB 2206 TTTTACTACTAATAATATATAAAA 2184

RESULT 7
US-08-948-265-3/c
; Sequence 3, Application US/08948265
; Patent No. 6228619
; GENERAL INFORMATION:
; APPLICANT: Foster, Simon J.
; APPLICANT: Burnham, Martin K. R.
; TITLE OF INVENTION: No. 6228619e1 glucosaminidase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,265

```

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,265
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/899008
; FILING DATE: 23-JUL-1996
; APPLICATION NUMBER: US 60/011888
; FILING DATE: 20-FEB-1996
; APPLICATION NUMBER: PCTUS97/02547
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Q. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-06
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 786 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-948-265-3

Query Match 3.8%; Score 38; DB 3; Length 786;
Best Local Similarity 49.0%; Pred. No. 0.26;
Matches 101; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 807 TGGGAGCTTTGTATCTCTGAGAGACAATACATATGCTCTGATGTTGTAACTTAC 866
DB 229 TAGCAATTTCTATTATTTTCTTCAACAAATTTACCATCTCTCTTTTGTCTGTACA 170

QY 867 TACCAAAACCTTAAAGATTCGCTTATTTCTGTTCTATTGATATGATCATCTACTGG 926
DB 169 TACCTTTCCCTTGTGATTTGAACACGCTTCATCAAGAGATTAATTCACATCAITTTAA 110

QY 927 TAAATCAAGTTCTTTCTAATAATGTAGAAGATCAGAAAATCCATAGAAGATATCAACA 986
DB 109 ACAATTTAGTTTCTTACCGATGATTAATACAGCAAAACAACTAAATAACTATCAATA 50

QY 987 TTTGAGTTCTATGTAATAAAAAA 1012
DB 49 GTAGCGTTCAATCGGTAACTTGAAA 24

RESULT 8
US-08-948-265-1/c
; Sequence 1, Application US/08948265
; Patent No. 6228619
; GENERAL INFORMATION:
; APPLICANT: Foster, Simon J.
; APPLICANT: Burnham, Martin K. R.
; TITLE OF INVENTION: No. 6228619e1 glucosaminidase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/948,265

```

; FILING DATE: 514
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/899008
; FILING DATE: 23-JUL-1996
; APPLICATION NUMBER: US 60/011888
; FILING DATE: 20-FEB-1996
; APPLICATION NUMBER: PCTUS97/02547
; FILING DATE: 19-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Q. Todd
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50444-06
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215/994-2252
; TELEFAX: 215/994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 789 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-948-265-1

Query Match 3.8%; Score 38; DB 3; Length 789;
Best Local Similarity 49.0%; Pred. No. 0.267; Indels 0; Gaps 0;
Matches 101; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 807 TGGAGCTTTGTTATCTTCCTGAGAGCAATACATACATGCTCTCTGATGTTGTAACTTTAC 866
Db 229 TAGCAATTTCAATTAATTTGTTCAACAATTTACCATCCTCTCTTTGTCGTACAA 170
QY 867 TACCAAAACCTATAAGATTGCTTATTTGTTCTATTGGATATGTCATCATCTACTGG 926
Db 169 TACCTTTCCTTCTGTCATTTGAACAGCCTCATCAAGAGTATTTACATCATTTTAA 110
QY 927 TAAATCAATTTCTTTCTTAATTAATAGAGATCAGAAAATCCATAAGAGATATCAACA 986
Db 109 ACAATTTAGTTTCATTCAGCATGAGTAAATACAGCAAAACCAACTAAATTAATCAATA 50
QY 987 TTGAGTTCTATGGTAAAAA 1012
Db 49 GTAGCTTGAATGCGTACTTGAAA 24

RESULT 9
US-09-489-847-89
; Sequence 89, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 1342
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-489-847-89
Query Match 3.8%; Score 38; DB 4; Length 1342;
Best Local Similarity 54.2%; Pred. No. 0.32;
Matches 77; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 871 AAAACCTATAAGATTGCTTATTTGTTCTATTGGATATGTCATCATCTACTGTTAAA 930
Db 1183 AAAGCTACCAATCTTTGTACAATTTGTAATGTTAAGATTTTATTATCTGTTAAA 1242
QY 931 TCAAGTTTCTTTCTTAATAATGTAAGATCAGAAAATCCATAAGAGATATCAACATTG 990
Db 1243 TAAAAAATTTTCCAAAAA 1302
QY 991 AGTTCTATGGTAAAAA 1012
Db 1303 AAAAAA 1324

RESULT 10
US-09-007-484-3
; Sequence 3, Application US/09007484
; Patent No. 6072032
; GENERAL INFORMATION:
; APPLICANT: Black, Michael T.
; TITLE OF INVENTION: No. 6072032el Ptsy
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/007,484
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10081
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-007-484-3

Query Match 3.7%; Score 37.8; DB 3; Length 831;
Best Local Similarity 51.5%; Pred. No. 0.3;
Matches 87; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 513 AGCTGGCTATGTTGCTTCCACAGTGAAGAACTCTCTAGGAGGTATCTTACACGCAT 572
Db 660 AACGGGAATTTGTTGACTGAATGGAATGCTCGAGGAGGTGTTCTTAGCCAT 719
QY 573 TCAGAAATTCATATCATCACCCCTGAGTGGTTTGCAAAAGAGATGCAATCGAAT 632

Db 720 TCGTGAAGAACTCAATATTCCTGTAAATTTGTTGGTTGGTGAAGAAATCGATGATAT 779
QY 633 AGGTATATGCTTCTATAAGATTTTAAGCCAGGAGTGTGTGGTTCA 681
Db 780 TGGAGAGTTTAACTCAGAAACTTTTATGAAGGTTCTTTGGAAGGCTTA 828

RESULT 11
US-09-309-682-3
; Sequence 3, Application US/09309682
; Patent No. 6214348
; GENERAL INFORMATION:
; APPLICANT: Black, Michael T.
; TITLE OF INVENTION: No. 6214348el Ftsy
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch St
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/309,682
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/007,484
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10081
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 831 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-309-682-3
Query Match 3.7%; Score 37.8; DB 3; Length 831;
Best Local Similarity 51.5%; Pred. No. 0.3; Mismatches 0; Gaps 0;
Matches 87; Conservative 0; Indels 0; Gaps 0;

QY 513 AGCTGGCTATGTTGCTTCCACAGTGAAGATTCCTCTAGGAGGTATCGTAACGCCAT 572
Db 660 AACGGGAATGTTTGTACTAAGATTGATGGAACCTGCTCGAGGAGGTGTTGTTCTAGCCAT 719
QY 573 TCAGATGCATTTTCATATCACTCCCTGAAGTGGTTTGCAGAGAGATGCAATCGATGAAT 632
Db 720 TCGTGAAGAACTCAATATTCCTGTAAATTTGTTGGTTGGTGAAGAAATCGATGATAT 779
QY 633 ACGTATATGCTTCTATAAGATTTTAAGCCAGGAGTGTGTGGTTCA 681
Db 780 TGGAGAGTTTAACTCAGAAACTTTTATGAAGGTTCTTTGGAAGGCTTA 828

RESULT 12
US-09-007-484-1
; Sequence 1, Application US/09007484
; Patent No. 6072032
; GENERAL INFORMATION:
; APPLICANT: Black, Michael T.

; TITLE OF INVENTION: No. 6072032el Ftsy
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch St
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/007,484
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10081
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1293 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-007-484-1
Query Match 3.7%; Score 37.8; DB 3; Length 1293;
Best Local Similarity 51.5%; Pred. No. 0.36; Mismatches 87; Conservative 0; Indels 0; Gaps 0;

QY 513 AGCTGGCTATGTTGCTTCCACAGTGAAGATTCCTCTAGGAGGTATCGTAACGCCAT 572
Db 1119 AACGGGAATGTTTGTACTAAGATTGATGGAACCTGCTCGAGGAGGTGTTGTTCTAGCCAT 1178
QY 573 TCAGATGCATTTTCATATCACTCCCTGAAGTGGTTTGCAGAGAGATGCAATCGATGAAT 632
Db 1179 TCGTGAAGAACTCAATATTCCTGTAAATTTGTTGGTTGGTGAAGAAATCGATGATAT 1238
QY 633 ACGTATATGCTTCTATAAGATTTTAAGCCAGGAGTGTGTGGTTCA 681
Db 1239 TGGAGAGTTTAACTCAGAAACTTTTATGAAGGTTCTTTGGAAGGCTTA 1287

RESULT 13
US-09-309-682-1
; Sequence 1, Application US/09309682
; Patent No. 6214348
; GENERAL INFORMATION:
; APPLICANT: Black, Michael T.
; TITLE OF INVENTION: No. 6214348el Ftsy
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch St
; CITY: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/309,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/007,484
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM10081
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1293 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-309-682-1

Query Match 3.7%; Score 37.8; DB 3; Length 1293;
Best Local Similarity 51.5%; Pred. No. 0.36;
Matches 87; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 513 AGTCGCTATGTTCTTCAACAGTCAAAAGATCTCTAGAGGATGCTAACAGCCAT 572
Db 1119 AACGGGAATGTTTTCAGTAAGATGATGGAAGTCTCAGAGGATGCTTCTAGCCAT 1178

QY 573 TCAGATGCAATTCATATCACCCCTGAAGTGGTTTGCMAAAGAGATGCAATCGATGAAT 632
Db 1179 TCGTGAAGAACTCAATATCTCTGTAATAATGATTGGTTTGGTGAATAATCGATGATAT 1238

QY 533 ACCTATATGCTTCTATAAAGATTAAAGCCAGGACTGTTGGTTCA 681
Db 1239 TCGAGAGTTTAACTCAGAAAACCTTATGAAGTCTCTTGGAGGCTTA 1287

RESULT 14
US-09-411-449-1
; Sequence 1, Application US/09411449
; Patent No. 6524851
; GENERAL INFORMATION:
; APPLICANT: James Ellis
; TITLE OF INVENTION: HYBRID NUCLEIC ACID MOLECULES AND VECTORS INCLUDING
; TITLE OF INVENTION: beta-GLOBIN REGULATORY ELEMENTS
; FILE REFERENCE: 17860017
; CURRENT APPLICATION NUMBER: US/09/411,449
; CURRENT FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 2,246,005
; PRIOR FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-411-449-1

Query Match 3.7%; Score 37.8; DB 4; Length 3496;
Best Local Similarity 48.8%; Pred. No. 0.56;
Matches 102; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 768 GATGCCAACAGAAAGAGAGCTCTTTGAATCGGAAAGATGGGAGCTTTGTTATCTTCTG 827
Db 915 GATTACCCCATTTGATAGTACACCTTTGGGTTGTAAGTGACTTTTATTTATTTGTTATTT 974

QY 828 AGACAAATACATACATGCTCTGATCTTGTAACTTTACTACCAAAACCTATAAAGATTG 887
Db 975 TGACTGCATTAAAGAGTCTCTAGTTTTTATCTCTGTTTCCCAAAACCTATAAAGTAAC 1034

QY 888 GCTTATTTTCTGTTCTATTGGATATGATCATCATCTGTTAAATCAAGTTTCTTTCTAAT 947
Db 1035 TAATGCACAGACACATTTGTTGTTATTTCTATTTTATTTAGACATAATTTATTAGCATG 1094

QY 948 AATGTAGAAGATCAGAAAATCCATAGAA 976
Db 1095 CATGAGCAAAATTAAGAAAAACAACAACAA 1123

Search completed: January 21, 2004, 22:19:50
Job time : 71.455 secs

SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/309,682
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/007,484
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM10081
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1293 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-309-682-1

Query Match 3.7%; Score 37.8; DB 3; Length 1293;
Best Local Similarity 51.5%; Pred. No. 0.36;
Matches 87; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 513 AGTCGCTATGTTCTTCAACAGTCAAAAGATCTCTAGAGGATGCTAACAGCCAT 572
Db 1119 AACGGGAATGTTTTCAGTAAGATGATGGAAGTCTCAGAGGATGCTTCTAGCCAT 1178

QY 573 TCAGATGCAATTCATATCACCCCTGAAGTGGTTTGCMAAAGAGATGCAATCGATGAAT 632
Db 1179 TCGTGAAGAACTCAATATCTCTGTAATAATGATTGGTTTGGTGAATAATCGATGATAT 1238

QY 533 ACCTATATGCTTCTATAAAGATTAAAGCCAGGACTGTTGGTTCA 681
Db 1239 TCGAGAGTTTAACTCAGAAAACCTTATGAAGTCTCTTGGAGGCTTA 1287

RESULT 14
US-09-411-449-1
; Sequence 1, Application US/09411449
; Patent No. 6524851
; GENERAL INFORMATION:
; APPLICANT: James Ellis
; TITLE OF INVENTION: HYBRID NUCLEIC ACID MOLECULES AND VECTORS INCLUDING
; TITLE OF INVENTION: beta-GLOBIN REGULATORY ELEMENTS
; FILE REFERENCE: 17860017
; CURRENT APPLICATION NUMBER: US/09/411,449
; CURRENT FILING DATE: 1999-10-01
; PRIOR APPLICATION NUMBER: 2,246,005
; PRIOR FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-411-449-1

Query Match 3.7%; Score 37.8; DB 4; Length 3496;
Best Local Similarity 48.8%; Pred. No. 0.56;
Matches 102; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 768 GATGCCAACAGAAAGAGAGCTCTTTGAATCGGAAAGATGGGAGCTTTGTTATCTTCTG 827
Db 915 GATTACCCCATTTGATAGTACACCTTTGGGTTGTAAGTGACTTTTATTTATTTGTTATTT 974

QY 828 AGACAAATACATACATGCTCTGATCTTGTAACTTTACTACCAAAACCTATAAAGATTG 887
Db 975 TGACTGCATTAAAGAGTCTCTAGTTTTTATCTCTGTTTCCCAAAACCTATAAAGTAAC 1034

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 21, 2004, 17:11.45 ; Search time 340.967 Seconds
(without alignments)
10461.606 Million cell updates/sec

Title: US-09-936-869-5
Perfect score: 1012
Sequence: 1 atcgataaagtcattggc.....ttctatgtgtaaaaaaaaaa 1012

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	208.4	20.6	1145	12	US-10-310-154-178
2	83.2	8.2	266	10	US-09-878-574-5684
3	76.6	7.6	229	10	US-09-878-574-5737
4	63.8	6.3	224	10	US-09-878-574-8638
5	46.8	4.6	693	10	US-09-838-842A-1871
6	43.4	4.3	628	13	US-10-027-632-84237
7	43.4	4.3	628	13	US-10-027-632-296997
8	43.4	4.3	628	14	US-10-027-632-84237
9	43.4	4.3	628	14	US-10-027-632-296997
10	40.8	4.0	19513	13	US-10-204-708-39
11	40.4	4.0	10133	13	US-10-311-455-431
12	39.6	3.9	3673778	13	US-10-312-841-1
13	39.4	3.9	1278	9	US-09-815-242-9165
14	39.2	3.9	381	10	US-09-878-574-1390
15	39.2	3.9	387	10	US-09-878-574-953

ALIGNMENTS

RESULT 1

US-10-310-154-178
; Sequence 178, Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Aneeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Banu, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jinzhao
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shlnshieh
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luethy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madson, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKiel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, Manchikanti

16	39.2	3.9	598	13	US-10-027-632-199664	Sequence 199664,
17	39.2	3.9	598	13	US-10-027-632-199665	Sequence 199665,
18	39.2	3.9	598	14	US-10-027-632-199664	Sequence 199664,
19	39.2	3.9	598	14	US-10-027-632-199665	Sequence 199665,
20	39	3.9	517	13	US-10-027-632-179698	Sequence 179698,
21	39	3.9	517	14	US-10-027-632-179698	Sequence 179698,
22	39	3.9	7810	13	US-10-240-453-155	Sequence 155, App
23	39	3.9	7810	15	US-10-239-676-141	Sequence 141, App
24	38.6	3.8	597	13	US-09-814-353-4872	Sequence 4872, Ap
25	38.6	3.8	597	13	US-09-814-353-11169	Sequence 11169, A
26	38.4	3.8	5520	13	US-10-311-455-1785	Sequence 1785, Ap
27	38.4	3.8	6106	13	US-10-311-455-1446	Sequence 1446, Ap
28	38.4	3.8	6123	13	US-10-311-455-793	Sequence 793, App
29	38.4	3.8	6265	13	US-10-311-455-1390	Sequence 1390, Ap
30	38.2	3.8	3673778	13	US-10-312-841-1	Sequence 1, Appli
31	38	3.8	786	9	US-09-770-075-3	Sequence 3, Appli
32	38	3.8	786	9	US-09-769-997-3	Sequence 3, Appli
33	38	3.8	789	9	US-09-770-075-1	Sequence 1, Appli
34	38	3.8	789	9	US-09-769-997-1	Sequence 1, Appli
35	38	3.8	1406	13	US-10-027-632-124113	Sequence 124113,
36	38	3.8	1406	14	US-10-027-632-124113	Sequence 3, Appli
37	37.8	3.7	831	9	US-09-827-663-3	Sequence 9491, Ap
38	37.8	3.7	1290	9	US-09-845-242-9491	Sequence 9491, Ap
39	37.8	3.7	1293	9	US-09-827-663-1	Sequence 97574, A
40	37.8	3.7	1932	13	US-10-027-632-97574	Sequence 97574, A
41	37.8	3.7	1932	13	US-10-027-632-100117	Sequence 100117,
42	37.8	3.7	1932	14	US-10-027-632-97574	Sequence 97574, A
43	37.8	3.7	1932	14	US-10-027-632-100117	Sequence 100117,
44	37.8	3.7	15592	13	US-10-311-455-1300	Sequence 1300, Ap
45	37.8	3.7	73308	10	US-09-954-456-2276	Sequence 2276, Ap

APPLICANT: Parnell, Laurence D.
 APPLICANT: Start, William G.
 APPLICANT: Tennesen, Dan
 APPLICANT: Vidya, K.R.
 APPLICANT: Wang, Haiyun
 APPLICANT: Xin, Zhanqun
 APPLICANT: Xu, Nanfei
 APPLICANT: Yang, Chunzhi
 APPLICANT: Zeng, Xiaoping
 APPLICANT: Zhang, Qiang
 APPLICANT: Zhao, Yajuan
 APPLICANT: Zhou, Li
 TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
 FILE REFERENCE: 38-15(52796)B
 CURRENT APPLICATION NUMBER: US/10/310,154
 CURRENT FILING DATE: 2002-12-04
 PRIOR APPLICATION NUMBER: 60/337,358
 PRIOR FILING DATE: 2001-12-04
 NUMBER OF SEQ ID NOS: 736
 SEQ ID NO 178
 LENGTH: 1145
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (55)..(888)
 OTHER INFORMATION:
 US-10-310-154-178

Query Match 20.6%; Score 208.4; DB 12; Length 1145;
 Best Local Similarity 61.2%; Pred. No. 2.3e-49;
 Matches 354; Conservative 0; Mismatches 221; Indels 3; Gaps 1;

QY 98 CTCAGAGGAGTTCGATTATTTTCGCTCTCTATCTTCAATGCGCTGGAACCTATGCGCGT 157
 Db 158 CGACGGGAGTTGAGTACTTTCGCTCTCTCTCTCCAGTGGCGGACCATCTGGCGCT 217
 QY 158 GNACTCGGCATTTGCTCCAAAACGTTCTCGACAGGCTCGATGCTCCAACTCAAT 217
 Db 218 CCACCGGCGCTGCTGCGCCACCAACGCTGCTGCGG---CTGGAGCGGCTCCAGAGCT 274
 QY 218 TCACAAATTCATGGGTATGCGCTGACATATACGATGTTGCGGCTTCATGTTGTTATC 277
 Db 275 TCACGATCCACGGGTATGCGCGGACTACGACGAGGAGCTGCGGCTGCTGCGCGC 334
 QY 278 GATCTGACTTTAAGAGAGAGAGATTTCAAGTTGATGATGCTCTTGAGAACTACTGCG 337
 Db 335 GCACCAATTCGAGATGGAAGAATATGCGACCTGAAGAGGCTGCTTGACAACTACTGCG 394
 QY 338 CTAGTCTCAGTTGTGGTTCTCCATCATCATGCAATGTTGGGAAAGGTCATTTTGGGGCC 397
 Db 395 COTCCCTCTACTGCTCCAAATCTGGAACATGCTTCAGCGGCAAGGAGCTCTCTGGGCTC 454
 QY 398 ACGATGGGAGAAACATGGGACTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 457
 Db 455 ACGATGGGAGAAACATGGGACTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 514
 QY 458 TCCTTACCACACTTAATCTCTACTTGAAGCATATGTCACGATGTCCTTTATCAAGCTG 517
 Db 515 TCACCTTGGCTTGAACCT 574
 QY 518 GCTATGTTGCTTCCACAGTGAAGAATGCTCTAGGAGTATCGTAACAGGCAATTCAGA 577
 Db 575 GGATACAGGTTTCAAAATGGAAGAATGATGACGAGGATGTCATCGATACCATCAAC 634
 QY 578 ATGCAATTCATATCACCCCTGAAGTGGTTTGAAGAAGAGATGCAATCGATGAATACGTA 637
 Db 635 ATGCTTTTGGGGATCGCCACAAATTTGTTGAAGAGGGGTTCAATCGAAGAACTTAGCT 694
 QY 638 TATGCTCTTAAGAGATTTAAGCCCGAGGACTGTT 675
 Db 695 TATGCTTCGACAAAGAAATGAAGCCCTCTGATTCCTT 732

RESULT 2

US-09-878-574-5684
 Sequence 5684, Application US/09878574
 Patent No. US20020110548A1
 GENERAL INFORMATION:
 APPLICANT: Byrum, Joseph R.
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Thompson, Michael D.
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 TITLE OF INVENTION: Plants
 FILE REFERENCE: 38-21(15401)B
 CURRENT APPLICATION NUMBER: US/09/878,574
 CURRENT FILING DATE: 2001-12-21
 PRIOR APPLICATION NUMBER: 09/333,535
 PRIOR FILING DATE: 1999-06-14
 NUMBER OF SEQ ID NOS: 15775
 SEQ ID NO 5684
 LENGTH: 266
 TYPE: DNA
 ORGANISM: Glycine max
 OTHER INFORMATION: Clone ID: 701097250H1
 US-09-878-574-5684

Query Match 8.2%; Score 83.2; DB 10; Length 266;
 Best Local Similarity 73.2%; Pred. No. 1.5e-13;
 Matches 120; Conservative 0; Mismatches 43; Indels 1; Gaps 1;
 QY 99 TCACAGGAGTTCGATTATTTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 158
 Db 104 TCAACGAGAGTTGACTATTTTCGCTTTTCGCG-GTGCAATGCGCGGCTACTTCTGCAACG 162
 QY 159 AACTCGCGCATTTGCTCCAAAACGTTCTGCTCAGAGGCTCGATGCTCCCACTCAATT 218
 Db 163 CACCGCGAGTTGCTGCGCCACCAACAGGTTGCTGACAGGCTCCAAATTTTCCACAGTGT 222
 QY 219 CACAAATTCAGGTTATGCGCTGACTATACGATGTTGCTGCG 262
 Db 223 CACCATACATGAGCTTGGCTGACTATATGATGAACTGCG 266

RESULT 3

US-09-878-574-5737
 Sequence 5737, Application US/09878574
 Patent No. US20020110548A1
 GENERAL INFORMATION:
 APPLICANT: Byrum, Joseph R.
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Thompson, Michael D.
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 TITLE OF INVENTION: Plants
 FILE REFERENCE: 38-21(15401)B
 CURRENT APPLICATION NUMBER: US/09/878,574
 CURRENT FILING DATE: 2001-12-21
 PRIOR APPLICATION NUMBER: 09/333,535
 PRIOR FILING DATE: 1999-06-14
 NUMBER OF SEQ ID NOS: 15775
 SEQ ID NO 5737
 LENGTH: 229
 TYPE: DNA
 ORGANISM: Glycine max
 OTHER INFORMATION: Clone ID: 701097310H1
 US-09-878-574-5737

Query Match 7.6%; Score 76.6; DB 10; Length 229;
 Best Local Similarity 67.5%; Pred. No. 1.1e-11;
 Matches 137; Conservative 0; Mismatches 64; Indels 2; Gaps 2;
 QY 99 TCACAGGAGTTCGATTATTTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 157
 Db 27 TCAACGAGAGTTGACTATTTTCGCTTTTCGCTTCTGATGCGCGGCTACTTCTGCAAC 86
 QY 158 GAACTCGCCATTTGTTGCTCCAAAACGTTGCTGACAGGCTCCGATGCTCCAACTCAAT 217

Db 87 GCACCCGAGTTCCTGACCCACCAACGCTTGTGCGAAGAGTCCCAATTTTCCAAAGTGT 146
QY 218 TCACAATTGATGGTTATGCGCTGACTATAAAGATGGTTGCTGGCC-TTCATGTTGTTAT 276
Db 147 TCACCATACATGACTCTGCGCTGACTATTATGATAGACCTGGCGGCTGTTGCTCT 206
QY 277 CGATCTGACTTTAAAGAGAGGA 299
Db 207 GGATCTAGTTTCGATCCTAAGA 229

RESULT 4
US-09-878-574-8638
; Sequence 8638, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 8638
; LENGTH: 224
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701101468H1
US-09-878-574-8638

Query Match 6.3%; Score 63.8; DB 10; Length 224;
Best Local Similarity 69.9%; Pred. No. 5.6e-08;
Matches 86; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
QY 99 TCAGAGGAGTTCGATTTATTCGCTCTATCTTCAATGGCTGGAACTATTGCGGTG 158
Db 102 TCAAGAGAGTTTGACTATTTTCGCTTTCGCTGCAATGGCCGGTACTTACTGCAACG 161
QY 159 AACTCGCATTTGCTTCCAAAACGCTTCTGCTGAGAGGCTCGATGCTCCAACTCAAT 218
Db 162 CACCGCAGTTTCTGCCCATCATCGGTTGCTGAGAGGCTCCCAATTTTCCACAGTGT 221
QY 219 CAC 221
Db 222 CAC 224

RESULT 5
US-09-938-842A-1871
; Sequence 1871, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPL300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 1871
; LENGTH: 693
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1871
Query Match 4.6%; Score 46.8; DB 10; Length 693;
Best Local Similarity 56.0%; Pred. No. 0.0094;
Matches 131; Conservative 0; Mismatches 97; Indels 6; Gaps 2;
QY 386 CATTTTGGGCGCCACGAGTGGGAGAAACATGGGACTTGTCTCTCTCTGTTTTCATGATG 445
Db 341 CGTTTGGGAGCAGGAATGGGAGAGCATGTGTTCTCTCTGAATCGGTTAT---CGATC 397
QY 446 AGTATAATTACTTCTCTTACCACTTAACTCTCTTCTTGAAGCATAAATGTACGATGTC 505
Db 398 AACATGATATTTTCAAAACCGCTCTTAACTTAAACAGAGAAACCAATCTCTCTGGAGCTC 457
QY 506 TTTATCAAGCTGGCTATGTTGCTTCCAAACAGTAAAAAGTATCTCTAGGAGGTATCGTAA 565
Db 458 TAACCAAGCCGGGA---TTAATCCGATGGAAATCTTACTCTTTGGAGAGCATAAGAG 514
QY 566 CAGCCATTCAAGATGCATTTCATATCACCCCTGAGTGGTTTGCAAAAGAGATG 619
Db 515 ATTCGATAAAAGAGTCAATTGGTTTCACTCTCTGGTTGAGTGTAAACAGAGATG 568

RESULT 6
US-10-027-632-84237/C
; Sequence 84237, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84237
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-84237

Query Match 4.3%; Score 43.4; DB 13; Length 628;
Best Local Similarity 49.8%; Pred. No. 0.085;
Matches 110; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
QY 761 TTCTGAAGATGCCACAGAGAGAGAGCTTTTGAATCGGAAAGATGGAGCTTTGTTA 820
Db 375 TTGATAAAATACCAAGAGTACATTTTATTTCTAAATGAAAAGTATTCATATGTTTA 316
QY 821 TCTTCTGAGAGACAATACATATCTCTGATGTGTGAACCTTTTACTACCAAACTATA 880
Db 315 TCACCTGAAAAGTTTATCTAAGCAGCTGTTTGTCTTGGAAATTTCAATAACAGACACTGA 256
QY 881 AAGATTGGCTTATTTTCTTATTGATATGATCATCATCTACTGTTAAATCAAGTTTCT 940

Db 255 AACAAATGTCATAAGTGAATATATACAGGCGACATTAATGAATATCT 196
Qy 941 TTCTAATAATGTAGAGATCAGAAATCCATAGAGATAT 981
Db 195 TGATAATATTTAAATTCATGAAAATAGTATATTAAGTCAT 155

RESULT 7

US-10-027-632-296997/c
; Sequence 296997, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 296997
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-296997

Query Match 4.3%; Score 43.4; DB 13; Length 628;
Best Local Similarity 49.8%; Pred. No. 0.085;
Matches 110; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
Qy 761 TTCTGAAGATGCCAAGAGAGAGCTCTTTGAATCGGAAAGATGGAGCTTTGTTA 820
Db 375 TTGATAAATACCAAGGATACATTTTATATCTTAAATGAAAAGTATTCAAATGTTA 316
Qy 821 TCTTCTGAGAGACATACATCTCTCTGATGTGTAACTTTACTACCAAAACCTATA 880
Db 315 TCACCTGAAAAGTTATCTAAGCAGCTGTTCTTGGAAATTTCAATACAGAACCTGA 256
Qy 881 AAGATTGGCTTATTTCTGTTCTTATGGATATGATCATCATCTAGTAAATCAAGTTTCT 940
Db 255 AAACAATGTGCATAAAGTGCAATTAATACAGGCCACATTAATCAATGAATATCT 196
Qy 941 TTCTAATAATGTAGAGATCAGAAATCCATAAGAGATAT 981
Db 195 TGATAATATTTAAATTCATGAAAATAGTATATTAAGTCAT 155

RESULT 8

US-10-027-632-84237/c
; Sequence 84237, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84237
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-84237
Query Match 4.3%; Score 43.4; DB 14; Length 628;
Best Local Similarity 49.8%; Pred. No. 0.085;
Matches 110; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
Qy 761 TTCTGAAGATGCCAAGAGAGAGCTCTTTGAATCGGAAAGATGGAGCTTTGTTA 820
Db 375 TTGATAAATACCAAGGATACATTTTATATCTTAAATGAAAAGTATTCAAATGTTA 316
Qy 821 TCTTCTGAGAGACATACATCTCTCTGATGTGTAACTTTACTACCAAAACCTATA 880
Db 315 TCACCTGAAAAGTTATCTAAGCAGCTGTTCTTGGAAATTTCAATACAGAACCTGA 256
Qy 881 AAGATTGGCTTATTTCTGTTCTTATGGATATGATCATCATCTAGTAAATCAAGTTTCT 940
Db 255 AAACAATGTGCATAAAGTGCAATTAATACAGGCCACATTAATCAATGAATATCT 196
Qy 941 TTCTAATAATGTAGAGATCAGAAATCCATAAGAGATAT 981
Db 195 TGATAATATTTAAATTCATGAAAATAGTATATTAAGTCAT 155

RESULT 9

US-10-027-632-296997/c
; Sequence 296997, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 296997
; LENGTH: 628
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-296997

QY 751 GAAGCTATCGTCTGAAGATGCCAACAGAGAGAGCTCTTTGAATCGGAAGATGGG 810
DB 863063 GAAGGATCTTTTGAAGAGAGAGATAATAATAGGTTTGAAGCGGAAATATAGAG 863122
QY 811 AGCTTTGTTATCTCTCGAGAGACAATACATACATGTCCTGATGTTGTAACTTTACTACC 870
DB 863123 GTTGTTTATAGTAGGAGGAGTGGATATGGTTTATTTTATAGGGAATAGGATG 863182
QY 871 AAAACCTATAAGATGGCTTATTCGTTCTATTTGGATATGATCATCATCTGTTAA 930
DB 863183 GAGAGGGGAGAGTATAGAAATTTGTTGAATTTGTTTGGTTTATTTGGGATATAGGTTAA 863242
QY 931 TCAGTCTCTCTCTAATAATAGTAGAGATCAGAAATCCATAGAGATATCAACATTTG 990
DB 863243 ATAGTTTTTATAGTTATTTTATAGTATAGTATAGTATAGTATAGTATAGTATAGT 863302
QY 991 AGTTCTATGTAATAAAAAA 1012
DB 863303 TGATTTGTTATATATTTAA 863324

RESULT 13
US-09-815-242-9165
; Sequence 9165, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR FILING DATE: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR FILING DATE: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9165
; LENGTH: 1278
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
; NAME/KEY: CDS
; LOCATION: (1)...(1278)
US-09-815-242-9165

Query Match 3.9%; Score 39.4; DB 9; Length 1278;
Best Local Similarity 52.1%; Pred. No. 1.9;
Matches 88; Conservative 0; Mismatches 81; Indels 0; Gaps 0;
QY 513 AGCTGGCTATCTGCTTCCACAGTGAAGATATCTCTAGAGAGTATCGTAACAGCAT 572
DB 1104 AACGGGAATGTTTGTACTAGATGATGGAAGTCTCGAGAGGTGTGTTCTAGCCAT 1163
QY 573 TCAGATGATCTTATCATATCACCCTCGAGTGGTTTGGCAAGAGATGCAATCGATGAAT 632

DB 1164 TCGTAGAGACTCAATATTCCTCTAAATGATGTTTGGTGAAGAAATCGATGATAT 1223
QY 633 ACCTATATGCTTCTATAAAGATTTTAAAGCCAGGAGCTGTGTTGGTTCA 681
DB 1224 TGGAGAGTTTAACTCAGAAACTTTTATGAAGGTCTCTTGGAAAGGTTTA 1272
RESULT 14
US-09-878-574-1390
; Sequence 1390, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 1390
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-038-Q1-B1-F11
US-09-878-574-1390

Query Match 3.9%; Score 39.2; DB 10; Length 381;
Best Local Similarity 58.8%; Pred. No. 1;
Matches 87; Conservative 0; Mismatches 58; Indels 3; Gaps 1;
QY 206 CTCCAATCTCAATTCACAAATTCATGGTTATGGCTGATATAACGATGGTTCGGCCTT 265
DB 226 CTGCTGCTGATTTTCAGCATTCATGAGCTCTGGCCTAACTTCAAGGATGGCTCATGGCCT 285
QY 266 CATGTTGTTATCGATC---TGACTTTAAAGAGAGAGATTTCAACGTTGATGATGGTC 322
DB 286 CAAACTGTGACCCCTGACAGTGTCTTTGATAAATCTCAGATCTCAGACCTCTATAGCAACA 345
QY 323 TTGAGAGTACTGCGCTAGTCTCAGTTG 350
DB 346 TGGAGAGGATGGCCATCCCTGAGCTG 373

RESULT 15
US-09-878-574-953
; Sequence 953, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 953
; LENGTH: 387
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-043-Q1-B1-H8
US-09-878-574-953

Query Match 3.9%; Score 39.2; DB 10; Length 387;
Best Local Similarity 58.8%; Pred. No. 1;

Matches	87;	Conservative	0;	Mismatches	58;	Indels	3;	Gaps	1;
QY	206	CTC	AACTCAATTCACAA	TTTCATGGGTTATGGCCTGACTATACGATGGTTCTGTCGCCTT	265				
Db	226	CTG	TGCTGATTTTCAGCA	TTTCATGGGTTATGGCCTGACTATACGATGGTTCTGTCGCCTT	285				
QY	266	CAT	GTTGTTATCGATC---	TGACTTTAAAGAGAAGGAGATTTCACGTTGATGGATGGTC	322				
Db	286	CAA	ACTGTGACCCCTGAC	GTGCTTTTGATTAATCTCAGATCTCAGACCTCATTAGCAACA	345				
QY	323	TTG	AGAAGTACTGGCCTA	GTCTCAGTTG	350				
Db	346	TGG	AGAAGGATTGCCCA	TCCCTGAGCTG	373				

Search completed: January 22, 2004, 01:20:48
 Job time : 349.967 secs